STIC Search Report
Biotech-Chem Library

TO: Phillip Gambel
Location: 3e81 / 3c70
Wednesday, June 09, 2004
Art Unit: 1644
Phone: 272-0844
Serial Number: 09 / 891943

From: Jan Delaval
Location: Biotech-Chem Library
Rem 1A51
Phone: 272-2504

jan.delaval@uspto.gov

Search Notes

sira
Search and Information
Resources Administration

From:	Gambel, Phillip
Sent:	Monday, June 07, 2004 9:23 AM
To:	Delaval, Jan
Subject:	09/891,943 sequence search

jan

please perform a sequence and a sequence interference search for

09 / 891,943 (gallatin)

SEQ ID NO: 1

SEQ ID NO: 2

thanx

phillip gambel
art unit 1644
272-0844

1644 mailbox 3c70

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner # : _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

* *

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept to unify of the invention. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

* *

STAFF USE ONLY	Type of Search		Vendors and cost where applicable
Searcher _____ Jen	NA Sequence (#) ✓	STN _____
Searcher Phone #: ___ 2504	AA Sequence (#) ✓	Dialog _____
Searcher Location _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up ___ 6/7	Bibliographic _____	Dr.Link _____
Date Completed: ___ 6/9	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time _____	Fulltext _____	Sequence Systems ✓
Clerical Prep Time ___ 10	Patent Family _____	WWW/Internet _____
Online Time ___ +20	Other _____	Other (specify) _____

PTO 1595 (8/01)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: June 8, 2004, 05:52:06 ; Search time 22742 Seconds
(without alignment)
7101.229 Million cell updates/sec

Title: US-09-891-943-1
Perfect score: 3726
Sequence: 1 tgaccttcggcactgtgctt........agcataaatttcatatgct 3726

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3726	100.0	3726	6	AR044695	AR044695 Sequence
2	3726	100.0	3726	6	AR052342	AR052342 Sequence
3	3726	100.0	3726	6	AR055140	AR055140 Sequence
4	3726	100.0	3726	6	AR158011	AR158011 Sequence
5	3726	100.0	3726	6	I13528	I13528 Sequence 1
6	3726	100.0	3726	6	I15787	I15787 Sequence 1
7	3726	100.0	3726	6	I92491	I92491 Sequence 1
8	3726	100.0	3726	6	AR222661	AR222661 Sequence
9	3726	100.0	3726	6	AR399748	AR399748 Sequence
10	3726	100.0	3726	6	AX448965	AX448965 Sequence
11	3726	100.0	3726	6	BD070942	BD070942 Novel hum
12	3699	99.3	3785	6	AR044775	AR044775 Sequence
13	3699	99.3	3785	6	AR052422	AR052422 Sequence
14	3699	99.3	3785	6	AR055220	AR055220 Sequence
15	3699	99.3	3785	6	AR158091	AR158091 Sequence
16	3699	99.3	3785	6	I92571	I92571 Sequence 98
17	3699	99.3	3785	6	AR222741	AR222741 Sequence
18	3699	99.3	3785	6	AR399828	AR399828 Sequence
19	3699	99.3	3785	6	AX449062	AX449062 Sequence
20	3699	99.3	3785	6	BD071022	BD071022 Novel hum
21	3583	96.2	3956	6	AR044774	AR044774 Sequence
22	3583	96.2	3956	6	AR052421	AR052421 Sequence
23	3583	96.2	3956	6	AR055219	AR055219 Sequence
24	3583	96.2	3956	6	AR158090	AR158090 Sequence
25	3583	96.2	3956	6	I92570	I92570 Sequence 97
26	3583	96.2	3956	6	AR222740	AR222740 Sequence
27	3583	96.2	3956	6	AR399827	AR399827 Sequence
28	3583	96.2	3956	6	AX449061	AX449061 Sequence
29	3583	96.2	3956	6	BD071021	BD071021 Novel hum
30	3484.6	93.5	3486	9	HSU37028	U37028 Human leuko
31	2315.2	62.1	3803	6	AR044732	AR044732 Sequence
32	2315.2	62.1	3803	6	AR052379	AR052379 Sequence
33	2315.2	62.1	3803	6	AR055177	AR055177 Sequence
34	2315.2	62.1	3803	6	AR158048	AR158048 Sequence
35	2315.2	62.1	3803	6	I92528	I92528 Sequence 52
36	2315.2	62.1	3803	6	AR222698	AR222698 Sequence
37	2315.2	62.1	3803	6	AR399785	AR399785 Sequence
38	2315.2	62.1	3803	6	AX449016	AX449016 Sequence
39	2304.6	61.9	3558	10	BD070979	BD070979 Novel hum
40	2304.2	61.8	3597	6	AF021334	AF021334 Rattus no
41	2304.2	61.8	3597	6	AR044733	AR044733 Sequence
42	2304.2	61.8	3597	6	AR052380	AR052380 Sequence
43	2304.2	61.8	3597	6	AR055178	AR055178 Sequence
44	2304.2	61.8	3597	6	AR158049	AR158049 Sequence
45	2304.2	61.8	3597	6	I92529	I92529 Sequence 54

ALIGNMENTS

RESULT 1
AR044695
LOCUS   AR044695   3726 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION   Sequence 1 from patent US 5817515.
ACCESSION   AR044695
VERSION   AR044695.1 GI:5966160
KEYWORDS   Unknown.
SOURCE   Unknown.
  ORGANISM   Unclassified.
REFERENCE   1 (bases 1 to 3726)
  AUTHORS   Gallatin,W.Michael. and Van der Vieren,M.
  TITLE   Human B2 integrin alpha subunit antibodies
  JOURNAL   Patent: US 5817515-A 1 06-OCT-1998;
FEATURES   Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a

```
source          1..3726
                /organism="unknown"
                /mol_type="unassigned DNA"
ORIGIN

Query Match     100.0%;   Score 3726;   DB 6;   Length 3726;
Best Local Similarity   100.0%;   Pred. No.0;
Matches 3726;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;
```

```
Qy    1  TGACCTTCGGCACTGTGCTTCTTCTGAGTGTCCTGGCTTCTTATCATGGATTCAACCTGG   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  TGACCTTCGGCACTGTGCTTCTTCTGAGTGTCCTGGCTTCTTATCATGGATTCAACCTGG   60

Qy   61  ATGTGGAGGAGGCCTACGATCTTCCAGGAGGATGCAGGCGGCTTTGGGCAGAGCGTGGTGC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  ATGTGGAGGAGGCCTACGATCTTCCAGGAGGATGCAGGCGGCTTTGGGCAGAGCGTGGTGC  120

Qy  121  AGTTCGGTGATCTTCGACTCGTGTGGGGAGCACCCCTGAGGTGGTCGGGCCAACCAGA    180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  AGTTCGGTGATCTTCGACTCGTGTGGGGAGCACCCCTGAGGTGGTCGGGCCAACCAGA    180

Qy  181  CGGGACGGCGGTGTATGACTCGCGCAGCTGCCACCGGCACCCCATCCCGCCTGCACA     240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CGGGACGGCGGTGTATGACTCGCGCAGCTGCCACCGGCACCCCATCCCGCCTGCACA     240

Qy  241  TCCGCCCTGAGGGCCGTGAACATGTCCTTGGGCCTGACCCTGGCCAGCCTCCACCAACGGCT  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TCCGCCCTGAGGGCCGTGAACATGTCCTTGGGCCTGACCCTGGCCAGCCTCCACCAACGGCT  300

Qy  301  CCCGGCTCCTGGCCTGTGCCCGACCCTGCACAGAGTCTGTGGGGAGAACTCATACTTCAA   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  CCCGGCTCCTGGCCTGTGCCCGACCCTGCACAGAGTCTGTGGGGAGAACTCATACTTCAA   360

Qy  361  AGGGTTCCTGCCTCCTGCTCGCGCTGGGCTCGCGCTGGGGATCATCAGACAGTCCCGACGCCA  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  AGGGTTCCTGCCTCCTGCTCGCGCTGGGCTCGCGCTGGGGATCATCAGACAGTCCCGACGCCA  420

Qy  421  CGCCAGAGTGTCCACATCAAGAGATGGACATGTCTTCCTGATTGACGGCTCTGGAAGCA   480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  CGCCAGAGTGTCCACATCAAGAGATGGACATGTCTTCCTGATTGACGGCTCTGGAAGCA   480

Qy  481  TTGACCAAAATGACTTTAACCAGATGAAGGGCTTTGTCAAGCTGTCATGGGCCAGTTTG   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  TTGACCAAAATGACTTTAACCAGATGAAGGGCTTTGTCAAGCTGTCATGGGCCAGTTTG   540

Qy  541  AGGGCACTGACACCCTGTTTGCACTGAGTACTGCAGTACTCAAACCTCCTGAAGATCCACTTCA  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  AGGGCACTGACACCCTGTTTGCACTGAGTACTGCAGTACTCAAACCTCCTGAAGATCCACTTCA  600

Qy  601  CCTTCACCCAATTCCGGAACCCCGAGCCAGCCAGCAGCCGCTGGTGGATCCCATGTCCAAC   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  CCTTCACCCAATTCCGGAACCCCGAGCCAGCCAGCAGCCGCTGGTGGATCCCATGTCCAAC   660

Qy  661  TGAAAGGCCTGACGTTCACGGGCATCGGCCACGGGCATCCTGACAGTGGTGACAGCTATTCATC  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  TGAAAGGCCTGACGTTCACGGGCATCGGCCACGGGCATCCTGACAGTGGTGACAGCTATTTCATC  720

Qy  721  ATAAGAATGGGCCCGAAAAGTGCCAAGAGATCTCATTGTCATCACAGATGGGCAGA    780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  ATAAGAATGGGCCCGAAAAGTGCCAAGAGATCTCATTGTCATCACAGATGGGCAGA    780

Qy  781  AGTACAAAGACCCCCTGGAATACAGTGATGTGATGGTCATCCCCCAGGCAGAGAGGCTGGCATCA  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  AGTACAAAGACCCCCTGGAATACAGTGATGTGATGGTCATCCCCCAGGCAGAGAGGCTGGCATCA  840

Qy  841  TCCGCTACGCTATCGGGGTGGGACACGCGTTTCCAGGGACCCACTGCCAGGCAGGAGCTGA   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  TCCGCTACGCTATCGGGGTGGGACACGCGTTTCCAGGGACCCACTGCCAGGCAGGAGCTGA   900

Qy  901  ATACCATCAGCTCAGCCGCCTCCCGGAGGACCACCGGTTCAAGGTGGACAACTTTGCAGCCC  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  ATACCATCAGCTCAGCCGCCTCCCGGAGGACCACCGGTTCAAGGTGGACAACTTTGCAGCCC  960
```

```
Qy  961  TTGGCAGCATCCAGAAGCAGCTGCGAGGAGAAGATCTATGCAGTTGAGGAGAACCAGTCCA  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  TTGGCAGCATCCAGAAGCAGCTGCGAGGAGAAGATCTATGCAGTTGAGGAGAACCAGTCCA  1020

Qy 1021  GGGCAAGCAGCTCCTTCCAGCACGAGATGTCCAAGAGAGTGTCCAGCAGCTTCAGCACACCCTCACAA  1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  GGGCAAGCAGCTCCTTCCAGCACGAGATGTCCAAGAGAGTGTCCAGCAGCTTCAGCACACCCTCACAA  1080

Qy 1081  TGGATGGCCTCTTCTGGGGCTGTGGGGAGCTTTAGCTGGTCTCTGGAGGTGCCTTCCTGT  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  TGGATGGCCTCTTCTGGGGCTGTGGGGAGCTTTAGCTGGTCTCTGGAGGTGCCTTCCTGT  1140

Qy 1141  ATCCCCAAATATGAGCCCACCTTCATCAACAACATGTCTCAGGAGAATGTGACATGAGGG  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  ATCCCCAAATATGAGCCCACCTTCATCAACAACATGTCTCAGGAGAATGTGACATGAGGG  1200

Qy 1201  ACTCTTACCTGGTTACTCACCCGAGCTAGCCCTGGGAGCGGGGTACAGAACCTGGTCC   1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  ACTCTTACCTGGTTACTCACCCGAGCTAGCCCTGTGGAAGGGGGTACAGAACCTGGTCC   1260

Qy 1261  TGGGGGCCCCGTACCAGCATACCCGGGAAGGTCGTCATCTTCACCCAGGTGTCCAGGC    1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261  TGGGGGCCCCGTACCAGCATACCCGGGAAGGCGTCATCTTCACCCAGGTGTCCAGGC    1320

Qy 1321  AATGGAGGAGAAGGCCGAAGTCACAGGGACGCAGATCGGGCTCCTACTTCGGGGCCTCCC  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321  AATGGAGGAGAAGGCCGAAGTCACAGGGACGCAGATCGGGCTCCTCACTTCGGGGCCTCCC  1380

Qy 1381  TCTGCTCCGTGGATGTGGCAGCGATGGCAGCACCGACCTGATCCTCATTGGGGCCGAGA  1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381  TCTGCTCCGTGGATGTGGCAGCGATGGCAGCACCGACCTGATCCTCATTGGGGCCGAGA  1440

Qy 1441  ATTACTATGAGCAGACCCGAGGGGCCCAGGTGTCCGTGTGTCCCTTGCCCTAGGGGCAGA  1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441  ATTACTATGAGCAGACCCGAGGGGCCCAGGTGTCCGTGTGTCCCTTGCCCTAGGGGCAGA  1500

Qy 1501  GGGTGCAGTGGCAGTGTGACGCTGTTCTCCGTGGTGAGCAGGGCCACCCCTGGGGCCGCT  1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501  GGGTGCAGTGGCAGTGTGACGCTGTTCTCCGTGGTGAGCAGGGCCACCCCTGGGGCCGCT  1560

Qy 1561  TTGGGGCAGCCCTGACAGTGTTGTTGGGGGATGTGAATGAGAAGCTGATAGACGTGGCCA  1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561  TTGGGGCAGCCCTGACAGTGTTGTTGGGGGATGTGAATGAGAAGCTGATAGACGTGGCCA  1620

Qy 1621  TGGGGGCCCCGGGAGAGCGAAACCGGGGTGCTGTCTACCTGTTTCACGGAGCCTCAG    1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621  TGGGGGCCCCGGGAGAGCGAAACCGGGGTGCTGTCTACCTGTTTCACGGAGCCTCAG    1680

Qy 1681  AATCCGGCATCACGCCCTCCCACAGCCAGGATTGCCAGCCGATTGCCCAGCTCTCCCCAGGC  1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681  AATCCGGCATCACGCCCTCCCCACAGCCAGGATTGCCAGCCGATTGCCCAGCTCTCCCCAGGC  1740

Qy 1741  TGCAGTATTTTGGGCAGGCGCTGAGTGGGGTCAGGACCTCACCCAGGATGGACTGATGG  1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741  TGCAGTATTTTGGGCAGGCGCTGAGTGGGGTTCAGGACCTCACCCAGGATGGACTGATGG  1800

Qy 1801  ACCTGGCCGTGGGGGCCCGGGGCCCCGGGGCCCAGTGTCGTCCTGCTCCAGGAGTCTGCCGGTGCTGCTGAAAG  1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801  ACCTGGCCGTGGGGGCCCGGGGCCCCGGGGCCCAGTGTCGTCCTGCTCCAGGAGTCTGCCGGTGCTGCTGAAAG  1860

Qy 1861  TGGGGGTGGCCATGAGATTCAGCCTGTGGAGTGGCCAAGGCTGTGTGTACCGGTGCTGGG  1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861  TGGGGGTGGCCATGAGATTCAGCCTGTGGAGTGGCCAAGGCTGTGTGTACCGGTGCTGGG  1920

Qy 1921  AAGAGAAGCCCAGTGCCGCCTGGAAGCGCACCACCGTCTGTCTCACCATCCAGAAAA    1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921  AAGAGAAGCCCAGTGCCGCCTGGAAGCGCACCACCGTCTGTCTCACCATCCAGAAAA    1980

Qy 1981  GCTCACTGGACCAGCTAGGTGACCATCCAAAGCTCTGTCAGGTTTGATCTGGCACTGACC  2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981  GCTCACTGGACCAGCTAGGTGACCATCCAAAGCTCTGTCAGGTTTGATCTGGCACTGACC  2040

Qy 2041  CAGGTCGTCTGACTTCTCGTGCCATTTTCAATGAAACCCACTTTGACTGAA             2100
```

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      June  8, 2004, 05:52:06 ; Search time 13747 Seconds
                                        (without alignments)
                                  8093.874 Million cell updates/sec

Title:        US-09-891-943-1
Perfect score: 3726
Sequence:      1 tgaccttggcactgtgctt........agcataaatttccatatgct 3726

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched:     27513289 seqs, 1493109276 residues

Total number of hits satisfying chosen parameters:     55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    EST:*
              1:  em_estba:*
              2:  em_esthum:*
              3:  em_estin:*
              4:  em_estmu:*
              5:  em_estov:*
              6:  em_estpl:*
              7:  em_estro:*
              8:  em_htc:*
              9:  gb_est1:*
              10: gb_est2:*
              11: gb_htc:*
              12: gb_est3:*
              13: gb_est4:*
              14: gb_est5:*
              15: em_estfun:*
              16: em_estcm:*
              17: em_gss_hum:*
              18: em_gss_inv:*
              19: em_gss_pln:*
              20: em_gss_vrt:*
              21: em_gss_fun:*
              22: em_gss_mam:*
              23: em_gss_mus:*
              24: em_gss_pro:*
              25: em_gss_rod:*
              26: em_gss_phg:*
              27: em_gss_vrl:*
              28: gb_gssl:*
              29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	DB ID	Description
1	1638.8	44.0	4056	11	BC042840	BC042840 Mus muscu
2	1404.4	37.7	3684	11	AK089521	AK089521 Mus muscu
3	912	24.5	3019	11	AK040921	AK040921 Mus muscu
4	844.8	22.7	3326	11	AK039444	AK039444 Mus muscu

No.	Score		Length	DB	Accession	Accession	Description
c							
5	488.4	13.1	1067	13	BQ711865	BQ711865	AGENCOURT
6	428	11.5	438	10	BF871208	BF871208	CMO-ET012
7	391.6	11.5	918	13	BQ942456	BQ942456	AGENCOURT
8	384.6	10.3	572	12	BM366605	BM366605	BS3200280
9	366.4	9.8	529	12	BM364453	BM364453	BS3200010
10	356.4	9.6	659	13	BY739005	BY739005	BY739005
11	351.8	9.4	648	10	BB636941	BB636941	BB636941
12	351.2	9.4	545	12	BM366368	BM366368	BM366250
13	336.8	9.0	713	13	BY764281	BY764281	BY764281
14	326.2	8.8	1210	11	AK089167	AK089167	Mus muscu
15	319	8.6	658	10	BF135162	BF135162	601780209
16	318.2	8.5	939	12	BI911010	BI911010	603068730
17	311.4	8.4	753	10	BF160615	BF160615	601768505
18	302.6	8.1	493	13	BF655065	BF655065	279453 MA
19	301.2	8.1	908	13	BQ714983	BQ714983	AGENCOURT
20	300.4	8.1	819	10	BE571108	BE571108	601334465
21	299	8.0	682	14	CB522420	CB522420	UI-M-GH0-
22	295.4	7.9	491	10	BF601810	BF601810	266881 MA
23	293.2	7.9	761	12	BI685527	BI685527	603309003
24	292.8	7.9	389	12	BI679978	BI679978	457288 MA
25	279.2	7.5	984	13	BQ708826	BQ708826	AGENCOURT
26	277	7.4	540	9	AI604269	AI604269	mtilh06.y
27	273	7.3	546	9	AA178276	AA178276	mtilh06.r
28	267	7.2	538	9	AI286783	AI286783	ub96e01.r
29	264.2	7.1	666	13	BY749040	BY749040	BY749040
30	263	7.1	663	13	BY748444	BY748444	BY748444
31	263	7.1	689	13	BY749906	BY749906	BY749906
32	261.8	7.0	493	10	BF470350	BF470350	UI-M-BH3-
33	260.4	7.0	684	13	BY749554	BY749554	BY749554
34	260.2	7.0	386	10	AW141017	AW141017	EST291040
35	259.8	7.0	692	13	BY748124	BY748124	BY748124
36	259.4	7.0	700	13	BY750010	BY750010	BY750010
37	258.4	6.9	482	12	BI683165	BI683165	464433 MA
38	255.8	6.9	702	13	BY748616	BY748616	BY748616
39	254.6	6.8	678	13	BY749866	BY749866	BY749866
40	254.2	6.8	372	10	AW446046	AW446046	84222 MAR
41	253.8	6.8	805	12	BI645825	BI645825	603275281
42	252.2	6.8	376	12	BM146659	BM146659	TCAAP1E66
43	252.2	6.8	722	13	BX644572	BX644572	DKFZp781J
44	250.8	6.7	656	13	BY747704	BY747704	BY747704
45	248.4	6.7	803	12	BI107542	BI107542	602893888

ALIGNMENTS

```
RESULT 1
BC042840
BC042840
LOCUS       BC042840       4056 bp    mRNA    linear   HTC 16-JAN-2003
DEFINITION  Mus musculus, similar to integrin alpha M; clone IMAGE:3988275,
            mRNA.
ACCESSION   BC042840
VERSION     BC042840.1  GI:27768957
KEYWORDS    HTC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4056)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JAN-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
```

```
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 30 Row: g Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10946645
This clone has the following problem: frame shifted.

FEATURES        Location/Qualifiers
source          1..4056
                /organism="Mus musculus"
                /mol_type="mRNA"
                /strain="CZECH II"
                /db_xref="taxon:10090"
                /clone="IMAGE:3988275"
                /tissue_type="Mammary tumor metastatized to lung. Tumor
                arose spontaneously from a senescent normal mammary
                (clonal) outgrowth infected with the virus MMTV."
                /clone_lib="NCI_CGAP_Lu29"
                /lab_host="DH10B"
                /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match     44.0%;   Score 1638.8;   DB 11;   Length 4056;
Best Local Similarity  68.0%;   Pred. No. 0;
Matches 2363;   Conservative  0;   Mismatches 1072;   Indels  41;   Gaps   4;
```

```
Qy     3 ACCTTCGGCACTGTGCTTCTTCTGAGTGTCCTTCTTATCATGGATTCAACCTGGAT     62
           ||||  | | |  ||||||||| |||    |  |  |  ||||||||||| |||||
Db    63 ACCTTGGATAGCCTTTCTTCTGCGTGTTGGGGTTTGTTTCTTGGCTTCAACTTGGAT    122

Qy    63 GTGGAGAGCCTACGACTCTTCCAGGAGGATGCAGGCGGCCGGCTTTGGGCAGAGCCGTGGTGCAG  122
          |  |||| |||| ||| ||||||    | ||||  | ||    |    |||| | |||||| |||
Db   123 GCAGAGAAGCTGACACATTTTCACATGGACCGGTGCTGAGTGCTCGGACGGTGCTCCCAG        182

Qy   123 TTCGGTGGATCTGACTCGTGGTGGGGAGCACCCCTGGAGGTGGTGCGGGCCAACCAGACG       182
          |  | || |  | ||| ||||||| |||||| ||||    ||     | |||||| ||
Db   183 TATGATAGTTCCTGGGTGGTGGTTGGAGCACCAAAGGAAATAAAAGCCACTAACCAAATA       242

Qy   183 GGACGGCTGTATGACTGGCCAGCTGCCACCCGATGCCAGCCCATCCCGCTGCACATC          242
          || | ||| |  | |||||  ||||||||| ||||||||||||||||| |||| |||
Db   243 GGTGGCCTTCACAAATGTGGCCATCACACAGGCAACTGTGAGCCCATCTCCCTCCAGGTG       302

Qy   243 CGCCCTGAGCCGTGAACATGTCCTTGGGCCTGACCCTGGCAGCCTGGCAGCGGCTCC          302
          | ||| |||| |||||||||||| | ||||||||||||||||| ||||||  | |||
Db   303 CCCCCAGAGGCTGTGAACATGTCCCTGGCCCTGGCCTGTGCTGCTGCCTGGCCACCCTTCC      362

Qy   303 CGGCTCCTGGCCTGTGGCCTCGTGGTGGTCTGCTACTGCTGCACCACACAGTGTGCTCCAG     362
          ||| ||||||||||||||    ||| |  || || || ||| || ||| |||||||||||
Db   363 TGGCGTGTTGGCTGTGTGGTCCTGCTGGGCTGGGCGGCCTGGGGGAGAGAATATATCTTGACA    422

Qy   363 GGTTCCTGCCTCCGCTGGGCTGGCGTCGGCTGGGATCATCATTCAAGCAGAGCAGCCCAGAA    422
          ||  | ||  |  || || ||| |  |    |||| ||||||| |||  | |||
Db   423 CCAGAGTGTCCACATCAAGCAACATCGTCTTCCTGATTGACGGCTCTGGAAGCATT          480

Qy   423 GG--GCTGCTTTCTACTGAGTTCATCATTCAAGCAGAGCAGCCCAGAA                   482
          ||  | |||| | | | ||| ||||||||||| |  | | |||||||
Db   481 CAGGAGTGTCCAAAGCAAGCAAGACATCTGTGTTCCTGATTGATGGCTCGGGGTAGCATC      540

Qy   483 GACCAAAATGACTTTAACCAGATGAAGGGCTTTGTCCAGAGCTGTGTCATGGGCCAGTTTGAG    542
          | | | | |||||||    |||| ||   |||||| |||||||||||||| ||||||  |
Db   541 AGTTCCACAGATTTTGAAAAAATGCTGGACTTTGTTAAAGCTGTGATGAGCCAGGCTTCAG      600

Qy   543 GGCACTGACACCCTGTTGCACTGATGCAGTACTCAAACCTCCTGAAGATCCACTTCACC       602
          |||  | |||   || | | |||||||||| |||  |  | || | | || | | |||
Db   601 AGACCTAGCACACGGTTCTCCCTGATGCAGTTCTCTGATTACTTCCGAGTACAATTTTACT     660

Qy   603 TTCACCCAATTCCGGACCACCGGCCAGCAGCCTGGTGGATCCCATCGTCCAACTG           662
          |||  | || |  || | ||| ||| || | || |||| |||| | | || | |
Db   661 TTCAACAACTTCATCTCCACGTCAAGGCCTCAAGGCCTTTAAGTCTGTTGGATTCGTGAAGGCAGCTA  720

Qy   663 AAAGGCCTGACGTTCACGGCCACGGCATCCTGACAGTGGTGACAGTGGTGGACGAGTGATTTCATCAT  722
          |||||||||||||||| |||||  | |||||| | ||||||| | |||||| ||||||  |  || |
Db   721 AGAGGGTACACATACACAGCGTCGGCTATCAAGCATGTCATACAGCACTGTTCACCACC        780

Qy   723 AAGAATGGGGCCCGAAAAGTGCCAAGAGAGATCCCATTGTCATCACAGATGGGCAGAAAG        782
          |||| |||||| | || | || |||| | | || | |||||||||| ||||||||| |||
Db   781 CAAAGTGGAGCTCGGCCAAGATGCCACCAAGGTCCTCATTGTCATCACTGATGGGAGAAAA     840

Qy   783 TACAAAGACCCCCTGGAATACAGTGATGTCATCCCCAGCAGCAGAAGGCTGGCATCATC        842
          || ||||  | |  |||||||||||| || ||||||||| | | |||| |||||||||
Db   841 CAAGGGGACAACTTGAGTTATGATAGTGTCATCCCCATGCAGGCCCTGCAAGCATCATT      900

Qy   843 CGCTACGCTATCGGGGTGGGACACGCTTTCCAGGGACCCACTGCCAGGCAGGAGCTGGAAT    902
          ||||| ||  |||||| || ||| |  ||||| | || |||||||||| | |||   |||
Db   901 CGTTATGCAATTGGGGTAGGAAAGGCCCTTTACAAATGAACAATTCCAAGCAACTTTGAGCCCTT  960

Qy   903 ACCATCAGCTCAGGCGCTTCCGGCGACCACGTGTTCAAGGTGACAACTTTGATGCTTTG      962
          |||||  | || ||| | | || |||||| |||||||  || |||||||||||||||||
Db   961 GCCATTGCATCGATGCCTTCCCATGAATACGTATTCAGCGTGGAGAACTTTGATGCTTTG    1020

Qy   963 GGCAGCATCCAGAAGCAGCTGCAGGAGAGATCATGGATGATCAGTTGAGGGAACCCAGTTCCAGG  1022
          |  |||||||||| |||| | ||||||| |||| ||  |||| ||| || |    |||  ||| |
Db  1021 AAGGATATTGAGAATTCAGCTGAAGGAGGAGATCTTTGCCATTGAGGGCACAGACACACCA     1080

Qy  1023 GCAAGCAGCTCCTTCCAGCACGAGATGTCCAAGAGGCTTCAGCAGGAGGCTTCAGCATG     1082
          ||| |||| |  ||  ||||||||||||||| |||| ||||| ||||| |||| |||
Db  1081 AGCAGCAGTACTTTTGAATTGGATGTCCCAGGAGGTTCAGTGCTGTGTTTACACCT       1140

Qy  1083 GATGGCCTCTTCCTGGGGGTGTGGGAGGTTTAGCTGGTTCTGGAGGTGCCTTCTTCTGTAT    1142
          |||| || || ||||||| ||||||| ||||||||||| ||||||| |||   |||  |
Db  1141 GATGGACCAGTTCTGGGGCGTGTGGGGAAGCTTCAGCTGGTCTCTGGAGAATGGGGAC      1200

Qy  1143 CCCCAAATATGAGCCCCACCTTCATCAACAGTGTCAGGAGAATGTGGACATGAGGGGAC     1202
          ||| |||||||||  ||||||||||||||||||||||||||| | |||| ||| ||||
Db  1201 CCCTCAAATATGAGACCCCACCTTCATCAACAGTGTCAGGAGGAACGAGGAGGAC        1260

Qy  1203 TCTTACCTGGGTTACTCCACCGAGCTAGCCTGTGGAAGGGGTACAGAACCTGGTCCTG     1262
          | ||||||||||||||||||||||| | |  |||||||||||  ||| |||| ||||
Db  1261 GCTTACCTGGGTACTCCACCGAGGGGTCTTTTGGAAGGGGGTCCACAGCCTGATCCTG      1320

Qy  1263 GGGCCCCCGCTACCAGCATACCGGGAAGGCTGTCATCTTCACCCAGGTGTCCAGGCAA     1322
          ||||||| ||  |||||| || ||| |||| ||||||||| ||||||  ||||||||
Db  1321 GGGGCCCCTGCCCACCAGCAACACAGGGGAAGGTTGTCATCTTTACCCAGGAATCCAGGCAC    1380

Qy  1323 TGGAGAAGAGGCCGAAGTCACAGGGACGCAGACGATGCCAGCACCGACCGGCTCC        1382
          |||||  |  |||| ||||||| ||| | | ||||| || ||||| | | |  |
Db  1381 TGGAGGCCCAGTGAAGTCTGAAGTCACATGATAGAGACGATAGCGCAGCACTGAAGTGCCTCCCTC  1440

Qy  1383 TGCTCCGTGGATGTGACAGCGATGCCAGCCGCTGATCCTCATTGGGGCCCCCAT         1442
          || || |||||| |||| | |||| |||| | |||| |  |||||| |||||||
Db  1441 TGTTCTGTGGACATGGACATGTCAGTGTCAGTGGAGTGCCTAGGGGCTGGGGGAG        1500

Qy  1443 TACTATGAGACAGCCGAGGGGCCCAGGTGCTCGGTGTCCGTGTGTCCCTAGGGGTGGGGGAG   1502
          ||||||||| || || |||||| ||| ||||| |||| ||||||||| |||||| ||||||-
Db  1501 TACTATGAGCACACCCGAGGGGCAGGGCCCAGTGCAGTGTGCCCCATGGGGGCCACCCCCAT    1559

Qy  1503 GTCCAGTGGCAGTGTGACGTGTTCTCCGTGGTGACGGCTGTGAGCGAGGGCCGCCGCTTT    1562
          |  ||||||||| | |||| |  |  |  |||| |||||  || |  |||  |||||||||
Db  1560 --CAGGTGCATTGTGGGACCACCCCTCATGGGGACGGGGCCATCTTCTTGGGGGCCGCATC    1617

Qy  1563 GGGGCAGCCCTGACAGTGTTGGGGATGTGAATGAGACAGAAGCTGGCATT             1622
          ||||| ||| |||||||| | ||||  || |||| || ||| | |||||
Db  1618 GGGGGCGGCCTCTGACAGTGCTAGGGGACGTGAATGGGGACAGTGCTCGGCAGATGGGGTCTATT  1677

Qy  1623 GGGGCCCCGGAGGAGGAGGACCGGGGTGTGGGGTGTCTTACCGTTCACGGAGCCTCAGAA   1682
          |||||||| ||||||||||||  ||||||| | |||| ||||  | |||  ||||||| |
Db  1678 GGTGCACCGGGGAGGAGGAGGACAGAGGCGTGTCTCTACATATTTCATTGGAGCCTCAAGA    1737

Qy  1683 TCCGGCATCAGCCCCTCCAGGACCACCGGCCAGCAGCCTGGTGGATCCCATCGTCCAACTG    1742
```

OM nucleic - nucleic search, using sw model

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:   June 8, 2004, 05:52:06 ; Search time 2021 Seconds
                           (without alignments)
                           7832.155 Million cell updates/sec

Title:           US-09-891-943-1
Perfect score:   3726
Sequence:        1 tgaccttcggcactgtgctt.........agcataaatttccatatgct 3726
Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched:        3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:    6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_29Jan04:*
                 1:  geneseqn1980s:*
                 2:  geneseqn1990s:*
                 3:  geneseqn2000s:*
                 4:  geneseqn2001as:*
                 5:  geneseqn2001bs:*
                 6:  geneseqn2002s:*
                 7:  geneseqn2003as:*
                 8:  geneseqn2003bs:*
                 9:  geneseqn2003cs:*
                 10: geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3726	100.0	3726	2	AAQ91712	Aaq91712 Human bet
2	3726	100.0	3726	2	AAT79220	Aat79220 Human bet
3	3726	100.0	3726	2	AAV31540	Aav31540 Human bet
4	3726	100.0	3726	2	AAV35236	Aav35236 Human Bet
5	3726	100.0	3726	2	AAV67281	Aav67281 Human alp
6	3726	100.0	3726	2	AAV08453	Aav08453 Human alp
7	3726	100.0	3726	2	AAA60014	Aaa60014 Human alp
8	3726	100.0	3726	6	ABK82404	Abk82404 Human cDN
9	3699	99.3	3785	2	AAT79275	Aat79275 Human bet
10	3699	99.3	3785	2	AAV35316	Aav35316 Human Bet
11	3699	99.3	3785	2	AAV67299	Aav67299 Human alp
12	3699	99.3	3785	2	AAV08458	Aav08458 Human alp
13	3699	99.3	3785	6	AAA60094	Aaa60094 Human alp
14	3699	99.3	3785	6	ABK82484	Abk82484 Human Bet
15	3583	96.2	3956	2	AAT79274	Aat79274 Human bet
16	3583	96.2	3956	2	AAV35315	Aav35315 Human alp
17	3583	96.2	3956	2	AAV67298	Aav67298 Human alp
18	3583	96.2	3956	2	AAV08457	Aav08457 Human alp
19	3583	96.2	3956	3	AAA60093	Aaa60093 Human alp
20	3583	96.2	3956	6	ABK82483	Abk82483 Human Bet
21	2315.2	62.1	3803	2	AAT79256	Aat79256 Mouse bet
22	2315.2	62.1	3803	2	AAV31575	Aav31575 Mouse alp
23	2315.2	62.1	3803	2	AAV35273	Aav35273 Mouse bet
24	2315.2	62.1	3803	2	AAV63852	Aav63852 Mouse alp
25	2315.2	62.1	3803	2	AAV08516	Aav08516 Mouse alp
26	2315.2	62.1	3803	3	AAA60051	Aaa60051 Mouse alp
27	2315.2	62.1	3803	6	ABK82441	Abk82441 Mouse cDN
28	2313.6	61.8	3803	2	AAQ91714	Aaq91714 Mouse alp
29	2304.2	61.8	3597	2	AAT79257	Aat79257 Rat beta
30	2304.2	61.8	3597	2	AAV31576	Aav31576 Rat alpha
31	2304.2	61.8	3597	2	AAV35274	Aav35274 Rat beta-
32	2304.2	61.8	3597	2	AAV67280	Aav67280 Rat alpha
33	2304.2	61.8	3597	2	AAV08476	Aav08476 Rat alpha
34	2304.2	61.8	3597	3	AAA60052	Aaa60052 Rat alpha
35	2302.6	61.8	3597	6	ABK82442	Abk82442 Rat cDNA
36	2300.2	61.8	3597	2	AAQ91715	Aaq91715 Rat alpha
37	2300.2	61.7	3528	2	AAQ91740	Aaq91740 Rat alpha
38	2300.2	61.7	3528	2	AAT79243	Aat79243 Rat beta
39	2300.2	61.7	3528	2	AAV31563	Aav31563 Rat alpha
40	2300.2	61.7	3528	2	AAV63807	Aav63807 Rat alpha
41	2300.2	61.7	3528	2	AAV08471	Aav08471 Rat alpha
42	2300.2	61.7	3528	2	AAA60037	Aaa60037 Rat cDNA
43	2300.2	61.7	3528	6	ABK82427	Abk82427 Rat cDNA
44	2289	61.4	3526	2	AAV35259	Aav35259 Rat beta-
45	2210.8	59.3	3519	2	AAQ91713	Aaq91713 Mouse alp

ALIGNMENTS

```
RESULT 1
AAQ91712
ID   AAQ91712 standard; cDNA; 3726 BP.
XX
AC   AAQ91712;
XX
DT   28-DEC-1995 (first entry)
XX
DE   Human beta-2 integrin alpha-d cDNA.
XX
KW   Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis;
KW   inflammatory bowel disease; asthma; ss.
XX
OS   Homo sapiens.
XX
FH   Key            Location/Qualifiers
FT   CDS            3..3488
FT                  /*tag= a
XX
PN   WO9517412-A1.
XX
PD   29-JUN-1995.
XX
PF   21-DEC-1994;   94WO-US014832.
XX
PR   23-DEC-1993;   93US-00173497.
PR   05-AUG-1994;   94US-00286889.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM, Van Der Vieren M;
XX
DR   WPI; 1995-240603/31.
DR   P-PSDB; AAR78166.
XX
PT   Alpha sub-unit polypeptide of human beta 2 integrin - used to identify
PT   potential antiinflammatory agents, for the treatment of graft
PT   arteriosclerosis, inflammatory bowel disease, asthma, etc.
XX
PS   Claim 1; Page 82-87; 172pp; English.
XX
CC   A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha
CC   -TM1 was used to screen a human spleen cDNA library to identify clone
CC   19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
CC   manipulated to allow expression of recombinant alpha-d subunit in COS and
CC   CHO cells
```

```
XX
SQ   Sequence 3726 BP; 817 A; 1064 C; 1054 G; 791 T; 0 U; 0 Other;

Query Match            100.0%; Score 3726; DB 2; Length 3726;
Best Local Similarity  100.0%; Pred. No.0;
Matches 3726; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy      1 TGACCTTCGGCACTGTGCTTCTTCTGAGTGTCCTTCTTATCATGATTCAACCTGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGACCTTCGGCACTGTGCTTCTTCTGAGTGTCCTTCTTATCATGATTCAACCTGG 60

Qy     61 ATGTGGAGGAGCCTACGATCTTCAGGAGGATGCAGCCGCTTGGGCAGAGCGTGGTGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGTGGAGGAGCCTACGATCTTCAGGAGGATGCAGCCGCTTGGGCAGAGCGTGGTGC 120

Qy    121 AGTTCGGTGGATCTCGACTCGTGGTGGGAGCACCCCTGAGGTGGTGGCGGCCAACCAGA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGTTCGGTGGATCTCGACTCGTGGTGGGAGCACCCCTGAGGTGGTGGCGGCCAACCAGA 180

Qy    181 CGGGACGGCGTGTATGACTGGCGAGCTGCCACCGGACGTGCCAGCCCATCCCGCTGCACA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGGGACGGCGTGTATGACTGGCGAGCTGCCACCGGACGTGCCAGCCCATCCCGCTGCACA 240

Qy    241 TCCGCCCTGAGGCCGTGAACATGTCCTTGGGCTGGCAGCCTCCACCAACGGCT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCGCCCTGAGGCCGTGAACATGTCCTTGGGCTGGCAGCCTCCACCAACGGCT 300

Qy    301 CCCGGCTCCTGGCCTGTGGCCGACCCTGCACAGAGTCTGTGGGAGAACTCATACTCAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCCGGCTCCTGGCCTGTGGCCGACCCTGCACAGAGTCTGTGGGAGAACTCATACTCAA 360

Qy    361 AGGGTTCCTGCCTCCTGCGCTGGGCTCGCGCTGGGCTCGCGATCATCCAGACAGCGCCA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGGGTTCCTGCCTCCTGCGCTGGGCTCGCGCTGGGCTCGCGATCATCCAGACAGCGCCA 420

Qy    421 CGCCAGAGTGTCCACATCAAGAGATGGACATGTCGTCTTCCTGATTGACGGCTCTGGAAGCA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CGCCAGAGTGTCCACATCAAGAGATGGACATGTCGTCTTCCTGATTGACGGCTCTGGAAGCA 480

Qy    481 TTGACCAAAATGACTTAACCAGATGAAGGGCTTTGTCCAAGCTGTCATGGGCCAGTTTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTGACCAAAATGACTTAACCAGATGAAGGGCTTTGTCCAAGCTGTCATGGGCCAGTTTG 540

Qy    541 AGGGCACTGACACCCGTGTTGCACTGATGCAGTACTCAAACCTCCTGAAGATCCACTTCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGGGCACTGACACCCGTGTTGCACTGATGCAGTACTCAAACCTCCTGAAGATCCACTTCA 600

Qy    601 CCTTCACCCAATTCCGGACCACCCCTGGTGATCAGCGCCTGGTGATCCCATCGTCCAAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCTTCACCCAATTCCGGACCACCCCTGGTGATCAGCGCCTGGTGATCCCATCGTCCAAC 660

Qy    661 TGAAAGGCCTGACGTTCACGGCCACGGGCATCCTGACAGTTGGTGACACAGCTATTTCATC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGAAAGGCCTGACGTTCACGGCCACGGGCATCCTGACAGTTGGTGACACAGCTATTTCATC 720

Qy    721 ATAAGAATGGGCCCGAAAAAGTGCCAAGAAGATCCTCATTGTCATCACAGATGGGCAGA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ATAAGAATGGGCCCGAAAAAGTGCCAAGAAGATCCTCATTGTCATCACAGATGGGCAGA 780

Qy    781 AGTACAAAGACCCCCTGGAATACAGTGATGTCATCCCCCAGGCAGAGAAGGCTGGCATCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGTACAAAGACCCCCTGGAATACAGTGATGTCATCCCCCAGGCAGAGAAGGCTGGCATCA 840

Qy    841 TCCGCTACGCTATCGGGGTGGGACACGCTTTCCAGGGACCCACTGCCAGGCAGGAGCTGA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCCGCTACGCTATCGGGGTGGGACACGCTTTCCAGGGACCCACTGCCAGGCAGGAGCTGA 900

Qy    901 ATACCATCAGTCAGTCAGCACGGTGTTCAAGGTGACAACTTTGCAGCCC 960
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ATACCATCAGTCAGTCAGCACGGTGTTCAAGGTGACAACTTTGCAGCCC 960

Qy    961 TTGGCAGCATCCAGAGAAGCAGCTGCGAGGAGAAGATCTATGCAGTTGAGGGAACCCAGTCCA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTGGCAGCATCCAGAGAAGCAGCTGCGAGGAGAAGATCTATGCAGTTGAGGGAACCCAGTCCA 1020

Qy   1021 GGGCAAGCAGCTCCTTCCAGCACGACGAGATGTCCAAGAAGGTTCAGCACACGCCCTCACAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GGGCAAGCAGCTCCTTCCAGCACGACGAGATGTCCAAGAAGGTTCAGCACACGCCCTCACAA 1080

Qy   1081 TGGATGGCCTCTTCCTGGGGACTGTGGGAGCTTTAGCTGGTCTGGTCTGGAGGTGCCTTCCTGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TGGATGGCCTCTTCCTGGGGACTGTGGGAGCTTTAGCTGGTCTGGTCTGGAGGTGCCTTCCTGT 1140

Qy   1141 ATCCCCAAATATGAGCCCCACCTTCATCAACATGTCTCAGGAGAATGTGGACATGAGGG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCCCCAAATATGAGCCCCACCTTCATCAACATGTCTCAGGAGAATGTGGACATGAGGG 1200

Qy   1201 ACTCTTACCTGGGTTACTCCACCGAGCTAGCCCTGTGGAAGGGGTACAGAACCTCGTCC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACTCTTACCTGGGTTACTCCACCGAGCTAGCCCTGTGGAAGGGGTACAGAACCTCGTCC 1260

Qy   1261 TGGGGGCCCCCGCCTACCACAGCCGGAAGGCTGTCATCTTCACCCAGGTGTCCAGGC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGGGGGCCCCCGCCTACCACAGCCGGAAGGCTGTCATCTTCACCCAGGTGTCCAGGC 1320

Qy   1321 AATGGAGGAAGAGGCCGAAGTCACAGGACACGCAGATCGGCTCCTACTTCGGGGCCTCCC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AATGGAGGAAGAGGCCGAAGTCACAGGACACGCAGATCGGCTCCTACTTCGGGGCCTCCC 1380

Qy   1381 TCTGCTCCGTGGATGTGACACGGATGCGCAGCACCGACCTGATCCTGATCCTGATCCTTGCGTCCCCCC 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TCTGCTCCGTGGATGTGACACGGATGCGCAGCACCGACCTGATCCTGATCCTTGCGTCCCCCC 1440

Qy   1441 ATTACTATGAGCAGACACCCGAGGGGCCCAGGTGCGTGCCTTGCGTCCCTAGGGGGCAGA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ATTACTATGAGCAGACACCCGAGGGGCCCAGGTGCGTGCCTTGCGTCCCTAGGGGGCAGA 1500

Qy   1501 GGGTGCAGTGGCAGTGACGCTGTGTTCCGTGGTGAGCAGGGCCACCCCTGGGGCGCT 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GGGTGCAGTGGCAGTGACGCTGTGTTCCGTGGTGAGCAGGGCCACCCCTGGGGCGCT 1560

Qy   1561 TTGGGGCAGCCCTGACAGTGTTCGGGGATGGAATGAGGACAAGCTGATAGACGTGGCCA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTGGGGCAGCCCTGACAGTGTTCGGGGATGGAATGAGGACAAGCTGATAGACGTGGCCA 1620

Qy   1621 TTGGGGCCCCGGGGGAGAGCAGGAGAACCGGGGTGCTCTACCTGTTTCACGGAGCCTCAG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TTGGGGCCCCGGGGGAGAGCAGGAGAACCGGGGTGCTCTACCTGTTTCACGGAGCCTCAG 1680

Qy   1681 AATCCGGCATCAGCCCCTCCACAGCGCAGCGATTGCCAGCTCCCCAGCTCTCCCCCAGGC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 AATCCGGCATCAGCCCCTCCACAGCGCAGCGATTGCCAGCTCCCCAGCTCTCCCCCAGGC 1740

Qy   1741 TGCAGTATTTTGGGCAGGGCCTGAGTGGGGGCTCAGGGCTGGGGTCAGGACTTCACCGGAGATGATGG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TGCAGTATTTTGGGCAGGGCCTGAGTGGGGGCTCAGGGCTGGGGTCAGGACTTCACCGGAGATGATGG 1800

Qy   1801 ACCTGGCCGTGGGGCCCCGGGGCCCCGGGGCCCAGGTGCTCCTGCCTCCTGCGTCCGGGTGCTGCGGGAAAG 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ACCTGGCCGTGGGGCCCCGGGGCCCCGGGGCCCAGGTGCTCCTGCCTCCTGCGTCCGGGTGCTGCGGGAAAG 1860

Qy   1861 TGGGGGTGGCCATGAGATTCAGCCTGTGGCCAAGGCGTGTCGTGGAGGTGGCCAAGGCGTGTGCTGG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGGGGGTGGCCATGAGATTCAGCCTGTGGCCAAGGCGTGTCGTGGAGGTGGCCAAGGCGTGTGCTGG 1920

Qy   1921 AAGAGAAGCCCAGTGCCCTGGAAGCTGGGGACGCCACCGTCTGTCTCCACCATCCAGAAA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AAGAGAAGCCCAGTGCCCTGGAAGCTGGGGACGCCACCGTCTGTCTCCACCATCCAGAAA 1980

Qy   1981 GCTCACTGGACCAGCTAGGTGACATCCAAAGCTCGTCAGGTTTGATCGGCACTGACC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GCTCACTGGACCAGCTAGGTGACATCCAAAGCTCGTCAGGTTTGATCGGCACTGACC 2040

Qy   2041 CAGGTCGTCGACTTCTCGTCGCCATTTCAATGAACCAAGAACCCCACTTTGACTCGAA 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 CAGGTCGTCGACTTCTCGTCGCCATTTCAATGAACCAAGAACCCCACTTTGACTCGAA 2100
```

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2004, 16:01:04 ; Search time 4734 Seconds
                                           (without alignments)
                                           3590.620 Million cell updates/sec

Title: US-09-891-943-1
Perfect score: 3726
Sequence: 1 tgacctcggcactgtgctt.........agcataaatttcatatgct 3726

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 295936 seqs, 228099010 residues

Total number of hits satisfying chosen parameters: 5591872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications NA:*
                 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
                14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
                17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3726	100.0	3726	9	US-09-350-259-1	Sequence 1, Appli
2	3726	100.0	3726	10	US-09-891-943-1	Sequence 1, Appli
3	3699	99.3	3785	9	US-09-350-259-98	Sequence 98, Appl
4	3699	99.3	3785	10	US-09-891-943-98	Sequence 98, Appl
5	3583	96.2	3956	9	US-09-350-259-97	Sequence 97, Appl
6	3583	96.2	3956	10	US-09-891-943-97	Sequence 97, Appl
7	2315.2	62.1	3803	9	US-09-350-259-52	Sequence 52, Appl
8	2315.2	62.1	3803	10	US-09-891-943-52	Sequence 52, Appl
9	2304.2	61.8	3597	9	US-09-350-259-54	Sequence 54, Appl
10	2304.2	61.8	3597	10	US-09-891-943-54	Sequence 54, Appl
11	2300.2	61.7	3528	9	US-09-350-259-36	Sequence 36, Appl
12	2300.2	61.7	3528	10	US-09-891-943-36	Sequence 36, Appl
13	2210.8	59.3	3519	9	US-09-350-259-45	Sequence 45, Appl
14	2210.8	59.3	3519	10	US-09-891-943-45	Sequence 45, Appl

15	2138	57.4	2499	9	US-09-350-259-96	Sequence 96, Appl
16	2138	57.4	2499	10	US-09-891-943-96	Sequence 96, Appl
17	1929.8	51.8	4704	17	US-10-641-643-1358	Sequence 1358, Ap
18	1628.8	43.7	4740	9	US-09-945-265-3	Sequence 3, Appli
19	1628.8	43.7	4740	10	US-09-902-481A-2	Sequence 2, Appli
20	1628.8	43.7	4740	15	US-10-207-655-175	Sequence 175, App
21	1628.8	43.7	4740	15	US-10-136-801-18	Sequence 18, Appl
22	1621	43.5	4124	17	US-10-641-643-1178	Sequence 1178, Ap
23	1611.2	43.2	3595	17	US-10-641-643-1147	Sequence 1147, Ap
24	1611.2	43.2	4740	16	US-10-159-563-339	Sequence 339, App
25	937	25.1	1484	9	US-09-350-259-102	Sequence 102, App
26	937	25.1	1484	10	US-09-891-943-102	Sequence 100, App
27	824.4	22.1	1318	9	US-09-350-259-100	Sequence 100, App
28	824.4	22.1	1318	10	US-09-891-943-100	Sequence 100, App
29	624.6	16.8	1006	9	US-09-350-259-24	Sequence 24, Appl
30	624.6	16.8	1006	10	US-09-891-943-24	Sequence 24, Appl
31	582.4	15.6	1586	15	US-10-240-965-4	Sequence 4, Appli
32	560.8	15.1	1891	16	US-10-108-260A-1724	Sequence 1724, Ap
33	494.2	13.3	852	9	US-09-350-259-92	Sequence 92, Appl
34	494.2	13.3	852	10	US-09-891-943-92	Sequence 92, Appl
35	477	12.8	1186	13	US-10-087-192-1211	Sequence 1211, Ap
36	297.6	8.0	4045	9	US-09-809-517A-37	Sequence 37, Appl
37	280.2	7.5	597	15	US-10-066-551-12	Sequence 12, Appl
38	279.2	7.5	750	13	US-09-925-298-35	Sequence 35, Appl
39	279.2	7.5	750	15	US-10-102-806-35	Sequence 35, Appl
40	279.2	7.5	750	16	US-10-264-049-21	Sequence 21, Appl
41	275	7.4	481	10	US-09-918-995-6255	Sequence 6255, Ap
42	268.2	7.2	670	15	US-10-002-631C-196	Sequence 196, App
43	256.4	6.9	5133	9	US-09-945-265-1	Sequence 1, Appli
44	256.4	6.9	5133	13	US-10-261-164-3	Sequence 3, Appli
45	256.4	6.9	5133	15	US-10-136-819-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-350-259-1
; Sequence 1, Application US/09350259
; Patent No. US2002006200BA1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: NO. US2002006200BA1e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3726
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(3485)
US-09-350-259-1

Query Match   100.0%; Score 3726; DB 9; Length 3726;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 3726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TGACCTCGGCACTGTGCTTCTTCTTCTGAGTGTCCTGGCTTCTTATCATGGATTCAACCTGG 60

Wed Jun 9 05:58:03 2004

us-09-891-943-1.rnpb

```
Db   1081  TGGATGGCCTCTTCCTGGGGGCTGTGGGGGCTGTGGGGGCTTTAGTCTGGTCTGGAGGGTGCCTTCCTGT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1141  ATCCCCCAAATATGAGCGACCTTCATCAACAACATGTCTCAGGAGGAATGTGACATGAGGG  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  ATCCCCCAAATATGAGCCCACCTTCATCAACAACATGTCTCAGGAGGAATGTGACATGAGGG  1200

Qy   1201  ACTCTTACCTGGTTACTCCACCGAGCTAGCCCTGTGGAAGGGGGTACAGAACCTGGTCC  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  ACTCTTACCTGGTTACTCCACCGAGCTAGCCCTGTGGAAGGGGTACAGAACCTGGTCC  1260

Qy   1261  TGGGGCCCCCGCTACCAGCATACCGGGAAGGCTGTCATCTTCACCCAGGTGTCCAGGC  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  TGGGGCCCCCGCTACCAGCATACCGGGAAGGCTGTCATCTTCACCCAGGTGTCCAGGC  1320

Qy   1321  AATGGAGGAAGAGGCCGAAGTCACAGGGACGCAGATCGGCTTCCTACTTCGGGGCCTCCC  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  AATGGAGGAAGAGGCCGAAGTCACAGGGACGCAGATCGGCTTCCTACTTCGGGGCCTCCC  1380

Qy   1381  TCTGCTCCGTGGATGTGGACAGCAGCGATGGCGCACCGACCTGATCCTCATTGGGGCCCCC  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  TCTGCTCCGTGGATGTGGACAGCAGCGATGGCGCACCGACCTGATCCTCATTGGGGCCCCC  1440

Qy   1441  ATTACTATGAGCAGATACCCGAGGGGCCCAGGTGTCCGTGTCCCTGCCCTAGGGGGCAGA  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  ATTACTATGAGCAGATACCCGAGGGGCCCAGGTGTCCGTGTCCCTGCCCTAGGGGGCAGA  1500

Qy   1501  GGGTCAGTGGCAGTGTGACCGTGTTCCGTGGTGAGCAGGGCCACCCCTGGGGGCCGCT  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  GGGTCAGTGGCAGTGTGACCGTGTTCCGTGGTGAGCAGGGCCACCCCTGGGGGCCGCT  1560

Qy   1561  TTGGGGCAGCCCTGACAGTGTTGGGGATGTGAATAGGAGACAAGCTGATAGACGTGGCCA  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  TTGGGGCAGCCCTGACAGTGTTGGGGATGTGAATAGGAGACAAGCTGATAGACGTGGCCA  1620

Qy   1621  TTGGGGCCCCGGGAGCAGGAGAACCGGGGTGCTGTCTACCTGTTCACGGAGCCTCAG  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  TTGGGGCCCCGGGAGCAGGAGAACCGGGGTGCTGTCTACCTGTTCACGGAGCCTCAG  1680

Qy   1681  AATCCGGCATCAGCGCCCTCCAGCTGAGTGGGGGCTGAGGATGGACTGATGG  1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  AATCCGGCATCAGCGCCCTCCAGCTGAGTGGGGGCTGAGGATGGACTGATGG  1740

Qy   1741  TGCAGTATTTGGGCAGGCGCTGAGTGGGGCCCGGGGGTCAGGAGTCTTCGCCGGGTGCTGAAAG  1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741  TGCAGTATTTGGGCAGGCGCTGAGTGGGGCCCGGGGGTCAGGAGTCTTCGCCGGGTGCTGAAAG  1800

Qy   1801  ACCTGGCCGTGGGGGCCCGGGGGCCAGGTGCTTCAGGAGTCTTCGGCGGGTGCTGAAAG  1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801  ACCTGGCCGTGGGGGCCCGGGGGCCAGGTGCTTCAGGAGTCTTCGGCGGGTGCTGAAAG  1860

Qy   1861  TGGGGGTGGCCATGAGATTCAGCCCTGTGGAGGTGGCCAAGGCTGTGTGTACCGGTGCTGGG  1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861  TGGGGGTGGCCATGAGATTCAGCCCTGTGGAGGTGGCCAAGGCTGTGTGTACCGGTGCTGGG  1920

Qy   1921  AAGAGAAGCCCAGTGCCCTGGAAGCTGGGGACGCCACCGTCTGTCTCCATCCAGAAAA  1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921  AAGAGAAGCCCAGTGCCCTGGAAGCTGGGGACGCCACCGTCTGTCTCCATCCAGAAAA  1980

Qy   1981  GCTCACTGACACCAGCTAGGTGACATCCAAAGCTCTGTCAGGTTTGATCTGGACTGACC  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981  GCTCACTGACACCAGCTAGGTGACATCCAAAGCTCTGTCAGGTTTGATCTGGACTGACC  2040

Qy   2041  CAGGTCGTCGTCGACTTCTCGTGCATTTCAATGAAAACCAAGAACCCACTTTGACTCGAA  2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  CAGGTCGTCGTCGACTTCTCGTGCATTTCAATGAAAACCAAGAACCCACTTTGACTCGAA  2100

Qy   2101  GAAAAACCTGGACTGGGGGATTCACTGTGAAACCCTGAAGCTGCTTTTGCCAGATTGTG  2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  GAAAAACCTGGACTGGGGGATTCACTGTGAAACCCTGAAGCTGCTTTTGCCAGATTGTG  2160

Qy   2161  TGGAGGATGGTGGTGAGCCCCATCATTCTGCACCTCAACTTCTCACTGGTGAGAGAGCCCA  2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  TGGAGGATGGTGGTGAGCCCCATCATTCTCCACCTCAACTTCTCACTGGTGAGAGAGCCCA  2220
```

```
Db   1     TGGACTTCCGCACTGCTGTCTTCTGAGTGTCTGGCTTCTTATCATGGATTCAACCTGG  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61    ATGTGGAGGAGCAGCTTACGAGATCTTCCAGAGAGGATGCAGCGGCTTGGCAGAGCGTGTGTC  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61    ATGTGGAGGAGCAGCCTACGAGATCTTCCAGGAGGATGCAGCGGCTTGGCAGAGCGTGTGTGC  120

Qy   121   AGTTCGGTTGGATCGGACTCGTGCGGTGGGAGCACCCCTGGAGTGGTGGCGGCCAACCAGA  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121   AGTTCGGTTGGATCGGACTCGTGCGGTGGGAGCACCCCTGGAGTGGTGGCGGCCAACCAGA  180

Qy   181   CGGGACGGCTGTATGACGTGCTGCCACTGGCCTGTGCCAGCCGCATCCGCTGCACA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181   CGGGACGGCTGTATGACGTGCTGCCACTGGCCTGTGCCAGCCGCATCCGCTGCACA  240

Qy   241   TCCGCCCTGAGGCCGTGAACATGTCCTTGGGCCTGACCCTGGCAGCCTCCACCAACGGCT  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241   TCCGCCCTGAGGCCGTGAACATGTCCTTGGGCCTGACCCTGGCAGCCTCCACCAACGGCT  300

Qy   301   CCCGGCTCCTGGCCTGTGCCGACCCTGCACAGAGTCTGTGGGGAGAACTCATACTCAA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301   CCCGGCTCCTGGCCTGTGCCGACCCTGCACAGAGTCTGTGGGGAGAACTCATATCTCAA  360

Qy   361   AGGGTTCCTGCCTCGTCGGCGCTCGGGCTCGCGCTGGGGAGATCATCCAGACGAGCGCCA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361   AGGGTTCCTGCCTCGTCGGCGCTCGGGCTCGCGCTGGGGAGATCATCCAGACGAGCGCCA  420

Qy   421   CGCCAGAGTGTCCACATCAAGAGATGGACATCGTCTTCCTGATTGACGGCTCTGGAAGCA  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421   CGCCAGAGTGTCCACATCAAGAGATGGACATCGTCTTCCTGATTGACGGCTCTGGAAGCA  480

Qy   481   TTGACCAAAATGACTTAACCAGATGAAGGGCTTGTCAAGCTGTCATGGGCCAGTTTG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481   TTGACCAAAATGACTTAACCAGATGAAGGGCTTGTCAAGCTGTCATGGGCCAGTTTG  540

Qy   541   AGGGACTGACACCCTGTTTGCACTGATGCAGTACTCAAACCTCCTGAAGATCCACTTCA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541   AGGGACTGACACCCTGTTTGCACTGATGCAGTACTCAAACCTCCTGAAGATCCACTTCA  600

Qy   601   CCTTCACCCAATTCCGGACCGTTCACGCGCCACGGGCATCCTGACAGCTATTCATC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601   CCTTCACCCAATTCCGGACCGTTCACGCGCCACGGGCATCCTGACAGCTATTCATC  660

Qy   661   TGAAAGGCCTGACGTTCACGGGGACGCCGAAAAAGTGCCAAGAGATCCAAGAGATGGGACAGA  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661   TGAAAGGCCTGACGTTCACGGGACGCCGAAAAAGTGCCAAGAGATCCTAAGAGATGGGACAGA  720

Qy   721   ATAAGAATGGGGCCCGAAAAAGTGCCAAGAGATGTCATCATCTGTCATCACAGATGGGCAGA  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721   ATAAGAATGGGGCCCGAAAAAGTGCCAAGAGATGTCATCATCTGTCATCACAGATGGGCAGA  780

Qy   781   AGTACAAAGACCCCCCTGGAATACAGTGATGTCATCCCCCAGGCAGGAGAAGGCTGGCATCA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781   AGTACAAAGACCCCCCTGGAATACAGTGATGTCATCCCCCAGGCAGGAGAAGGCTGGCATCA  840

Qy   841   TCCGCTACGCTATCGGGGTGGGACACGCTTTCCAGGGACCCACTGCCAGGCAGGAGCTGA  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841   TCCGCTACGCTATCGGGGTGGGACACGCTTTCCAGGGACCCACTGCCAGGCAGGAGCTGA  900

Qy   901   ATACCATCAGCTCAGCGCCTCCGCAGAGACCACGTGTTCAAGGTGACAACTTTGCAGCCC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901   ATACCATCAGCTCAGCGCCTCCGCAGAGACCACGTGTTCAAGGTGACAACTTTGCAGCCC  960

Qy   961   TTGGCAGCATCCAGAGCAGCAGTGCAGGAGAAGATCTATGCAGTTGAGGGAGAACCCAGTCCA  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961   TTGGCAGCATCCAGAGCAGCAGTGCAGGAGAAGATCTATGCAGTTGAGGGAGAACCCAGTCCA  1020

Qy   1021  GGGCAAGCAGCTCCTTCCAGCACGAGATGTCCCAGGCAGAAGGCTTCAGCAGAGGCCCTCACAA  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  GGGCAAGCAGCTCCTTCCAGCACGAGATGTCCCAGGCAGAAGGCTTCAGCAGAGGCCCTCACAA  1080

Qy   1081  TGGATGGCCTCTTCCTGGGGGCTGTGGGGGCTTTAGTCTGGTCTGGAGGGTGCCTTCCTGT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   June 8, 2004, 05:52:06 ; Search time 282 Seconds
                                    (without alignments)
                                7332.436 Million cell updates/sec

Title:          US-09-891-943-1
Perfect score:  3726
Sequence:       1 tgacccttcggcactgtgctt.........agcataaatttcatatgct 3726

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:   1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued Patents NA:*
              1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
              2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
              3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
              4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
              5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
              6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3726	100.0	3726	1	US-08-173-497-1	Sequence 1, Appl
2	3726	100.0	3726	1	US-08-286-889-1	Sequence 1, Appl
3	3726	100.0	3726	1	US-08-485-618-1	Sequence 1, Appl
4	3726	100.0	3726	1	US-08-362-652-1	Sequence 1, Appl
5	3726	100.0	3726	1	US-08-605-672-1	Sequence 1, Appl
6	3726	100.0	3726	2	US-08-482-293A-1	Sequence 1, Appl
7	3726	100.0	3726	3	US-08-943-363-1	Sequence 1, Appl
8	3726	100.0	3726	3	US-09-193-043-1	Sequence 1, Appl
9	3726	100.0	3726	4	US-09-688-307A-1	Sequence 1, Appl
10	3726	100.0	3726	4	US-09-350-259-1	Sequence 1, Appl
11	3699	99.3	3785	1	US-08-485-618-98	Sequence 98, Appl
12	3699	99.3	3785	1	US-08-605-672-98	Sequence 98, Appl
13	3699	99.3	3785	2	US-08-482-293A-98	Sequence 98, Appl
14	3699	99.3	3785	2	US-08-943-363-98	Sequence 98, Appl
15	3699	99.3	3785	3	US-09-193-043-98	Sequence 98, Appl
16	3699	99.3	3785	4	US-09-688-307A-98	Sequence 98, Appl
17	3699	99.3	3785	4	US-09-350-259-98	Sequence 98, Appl
18	3583	96.2	3956	1	US-08-485-618-97	Sequence 97, Appl
19	3583	96.2	3956	1	US-08-605-672-97	Sequence 97, Appl
20	3583	96.2	3956	2	US-08-482-293A-97	Sequence 97, Appl
21	3583	96.2	3956	3	US-08-943-363-97	Sequence 97, Appl
22	3583	96.2	3956	3	US-09-193-043-97	Sequence 97, Appl
23	3583	96.2	3956	4	US-09-688-307A-97	Sequence 97, Appl
24	3583	96.2	3956	4	US-09-350-259-97	Sequence 97, Appl
25	2315.2	62.1	3803	1	US-08-485-618-52	Sequence 52, Appl
26	2315.2	62.1	3803	1	US-08-362-652-52	Sequence 52, Appl
27	2315.2	62.1	3803	1	US-08-605-672-52	Sequence 52, Appl
28	2315.2	62.1	3803	2	US-08-482-293A-52	Sequence 52, Appl
29	2315.2	62.1	3803	2	US-09-943-363-52	Sequence 52, Appl
30	2315.2	62.1	3803	3	US-09-193-043-52	Sequence 52, Appl
31	2315.2	62.1	3803	4	US-09-688-307A-52	Sequence 52, Appl
32	2315.2	62.1	3803	4	US-09-350-259-52	Sequence 52, Appl
33	2304.2	61.8	3597	1	US-08-485-618-54	Sequence 54, Appl
34	2304.2	61.8	3597	1	US-08-362-652-54	Sequence 54, Appl
35	2304.2	61.8	3597	1	US-08-605-672-54	Sequence 54, Appl
36	2304.2	61.8	3597	2	US-08-482-293A-54	Sequence 54, Appl
37	2304.2	61.8	3597	2	US-08-943-363-54	Sequence 54, Appl
38	2304.2	61.8	3597	3	US-09-193-043-54	Sequence 54, Appl
39	2304.2	61.8	3597	4	US-09-688-307A-54	Sequence 54, Appl
40	2304.2	61.8	3597	4	US-09-350-259-54	Sequence 54, Appl
41	2300.2	61.7	3528	1	US-08-286-889-36	Sequence 36, Appl
42	2300.2	61.7	3528	1	US-08-485-618-36	Sequence 36, Appl
43	2300.2	61.7	3528	1	US-08-362-652-36	Sequence 36, Appl
44	2300.2	61.7	3528	1	US-08-605-672-36	Sequence 36, Appl
45	2300.2	61.7	3528	2	US-08-482-293A-36	Sequence 36, Appl

ALIGNMENTS

```
RESULT 1
US-08-173-497-1
; Sequence 1, Application US/08173497
; Patent No. 5437958
;; GENERAL INFORMATION:
;; APPLICANT: Gallatin, W. Michael
;; APPLICANT: Van Der Vieren, Monica
;; TITLE OF INVENTION: No. 5437958el Human  2  Integrin Alpha
;; TITLE OF INVENTION: Subunit
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;; STREET: 233 S. Wacker Drive, 6300 Sears Tower
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60606-6402
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release 1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/173,497
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5437958and, Greta E.
;; REGISTRATION NUMBER: 35,302
;; REFERENCE/DOCKET NUMBER: 27866/31363
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312-474-6300
;; TELEFAX: 312-474-0448
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3726 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 3..3485
US-08-173-497-1

Query Match             100.0%;  Score 3726;  DB 1;  Length 3726;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 3726;  Mismatches    0;  Conservative   0;  Indels   0;  Gaps   0;
```

```
Qy     1  TGACCTTCGGCACTGCGTCTTCTCGAGTGTCCGGCGCTTCTTATCATGGATTCAAACCTGG  60
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     1  TGACCTTCGGCACTGCGTCTTCTGAGTGTCCGGCGCTTCTTATCATGGATTCAAACCTGG  60

Qy    61  ATGTGGAGGAGCCTACGATCTTCCAGGAGGATGCAGGCGCTTGGCCAGAGCGTGGTGC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  ATGTGGAGGAGCCTACGATCTTCCAGGAGGATGCAGGCGCTTGGCCAGAGCGTGGTGC  120

Qy   121  AGTTCGGTTGGATCTGACTCGTGTGGGAGCACCCCTGAGGTGGTGGCGGCCAACCAGA  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  AGTTCGGTTGGATCTGACTCGTGTGGGAGCACCCCTGAGGTGGTGGCGGCCAACCAGA  180

Qy   181  CGGGAGGCTGTATGACTGGCAGTGCCACGCATGTGCCAGCCCATCCCGCTGCACA  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  CGGGACGGCTGTATGACTGGCAGTGCCACGCATGTGCCAGCCCATCCCGCTGCACA  240

Qy   241  TCCGCCCTGAGGCCGTGAACATGTCCTTGGGCCTGGCAGCCTCCACCAACCGGCT  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TCCGCCCTGAGGCCGTGAACATGTCCTTGGGCCTGGCAGCCTCCACCAACCGGCT  300

Qy   301  CCCGGCTCCTGGCCTGCGTGTGGGCCCGCCTGCACAGAGTCGTGGGGAGAACTCATACTCAA  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  CCCGGCTCCTGGCCTGCGTGTGGGCCCGCCTGCACAGAGTCGTGGGGAGAACTCATACTCAA  360

Qy   361  AGGGTTCCTGCCTCCTGCTCGCGCTCGGGCTTCGCGCTGCGCTGGGAGATCATCCAGACAGTCCCCGACGCCA  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  AGGGTTCCTGCCTCCTGCTCGCGCTCGGGCTTCGCGCTGCGCTGGGAGATCATCCAGACAGTCCCCGACGCCA  420

Qy   421  CGCCAGAGTGTCCACATCAAGAAGGATGGATGGACATGGAACATGGAAGACCAGGCA  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  CGCCAGAGTGTCCACATCAAGAAGGATGGATGGACATGGAACATGGAAGACCAGGCA  480

Qy   481  TTGACCAAAATGACTTTAACCAGAGGGCCTTTGTCCAAGCTGTCATGGGCCAGTTTG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  TTGACCAAAATGACTTTAACCAGAGGGCCTTTGTCCAAGCTGTCATGGGCCAGTTTG  540

Qy   541  AGGGCACTGACACCCCTGTTTGCACTCGATGAGTACTCAAACCTCCTGAAGATCCACTTCA  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  AGGGCACTGACACCCCTGTTTGCACTGAGTACTCAAACCTCCTGAAGATCCACTTCA  600

Qy   601  CCTTCACCCAATTCCGGACCGCCTCCGGACCGCCGCAGAGCCCGAGCTGGTGGATCCCATCGTCCAAC  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CCTTCACCCAATTCCGGACCGCCTCCGGACCGCCGCAGAGCCCGAGCTGGTGGATCCCATCGTCCAAC  660

Qy   661  TGAAAGGCCTGACGTTCACGGCGCCTCACGCGCCTCGAGGAGGATCCTGACACAGCCTATTTCATC  720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  TGAAAGGCCTGACGTTCACGGCGCCTCACGCGCCTCGAGGAGGATCCTGACACAGCCTATTTCATC  720

Qy   721  ATAAGAATGGGGCCCCCTGGAATACAGTGATGTCATCCCCAAGAAGTCCAAGATCCTCATTGTCATCACAGATGGGCAGA  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  ATAAGAATGGGGCCCCCTGGAATACAGTGATGTCATCCCCAAGAAGTCCAAGATCCTCATTGTCATCACAGATGGGCAGA  780

Qy   781  AGTACAAAGAGAATCAGCTCCTTCCTTCCAGCACGATGTGTCCCCAGCAGGCTTCAGCACAGCCCTCACAA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  AGTACAAAGAGAATCAGCTCCTTCCTTCCAGCACGATGTGTCCCCAGCAGGCTTCAGCACAGCCCTCACAA  840

Qy   841  TCCGGTACGCTATCGGGGTGGGACACGCGCTTTCCAGGGACCCACTGCCAGGCAGGAGCTGA  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  TCCGGTACGCTATCGGGGTGGGGACACGCGCTTTCCAGGGACCCACTGCCAGGCAGGAGCTGA  900

Qy   901  ATACCATCAGCTCAGCGCCTCAGCGCCTCCGCAGGACCACGTGTTCAAGGTGGACAACTTTGCAGCCC  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  ATACCATCAGCTCAGCGCCTCAGCGCCTCCGCAGGACCACGTGTTCAAGGTGGACAACTTTGCAGCCC  960

Qy   961  TTGGCAGCATCCAGAGAAGCAGCTGCAGGAGAAGTCCCCAAGAAGGCTTGAGGGGAACCCAGTCCA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  TTGGCAGCATCCAGAGAAGCAGCTGCAGGAGAAGTCCCCAAGAAGGCTTGAGGGGAACCCAGTCCA  1020

Qy  1021  GGGCAAGCAGCAGCTCCTTCCAGCACACAGAGATGTCCCAGCACAGCCCTCACACAA  1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  GGGCAAGCAGCAGCTCCTTCCAGCACACAGAGATGTCCCAGCACAGCCCTCACACAA  1080
```

```
Qy  1081  TGGATGGCCCTCTTCCTGGGGGCTGTGGGAGGCTTTAGCTGGTCTCGGAGGTGCCTTCCTGT  1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TGGATGGCCCTCTTCCTGGGGGCTGTGGGAGGCTTTAGCTGGTCTCGGAGGTGCCTTCCTGT  1140

Qy  1141  ATCCCCAAATATGAGCCCCACCTTCATCAACATGTCTCAGGAGAATGTGACATGAGGG  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  ATCCCCAAATATGAGCCCCACCTTCATCAACATGTCTCAGGAGAATGTGACATGAGGG  1200

Qy  1201  ACTCTTACCTGGTTACTCACCGAGCTAGCCGTGTGGGAAGGGGTACAGAACCTGGTCC  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  ACTCTTACCTGGTTACTCACCGAGCTAGCCGTGTGGGAAGGGGTACAGAACCTGTCC  1260

Qy  1261  TGGGGGCCCCCGCTACCGGGAAGGACGCTGTCATCTTCACCCAGGTGTCCAGGC  1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  TGGGGGCCCCCGCTACCGGGAAGGACGCTGTCATCTTCACCCAGGTGTCCAGGC  1320

Qy  1321  AATGGAGGAAGGCCGAAGTCACAGGACGCAGATCGGCTCCTACTTCGGGGGCTCCC  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  AATGGAGGAAGGCCGAAGTCACAGGACGCAGATCGGCTCCTACTTCGGGGGCTCCC  1380

Qy  1381  TCTGCTCCGTGATGTGACAGCGCAGCGATGGCGAGCACCCGAGGGGCCAGGGGGGCAGA  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  TCTGCTCCGTGATGTGACAGCGCAGCGATGGCGAGCACCCGAGGGGCCAGGGGGGCAGA  1440

Qy  1441  ATTACTATGAGCAGACCCGAGGGGCCAGGTGTCCGTGTGCCTAGGGGGCAGA  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  ATTACTATGAGCAGACCCGAGGGGCCAGGTGTCCGTGTGCCTAGGGGGCAGA  1500

Qy  1501  GGGTGCAGTGGCAGTGTGACGGCTGTCTCCGTGGTGAGCGGCCACCCCTGGGGCCGCT  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  GGGTGCAGTGGCAGTGTGACGGCTGTCTCCGTGGTGAGCGGCCACCCCTGGGGCCGCT  1560

Qy  1561  TTGGGGCAGCCCCTGACACGTGTTGGGGGATGTGAATGACAGCAAGCTGTTCACCGGAGCGTGGCCA  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TTGGGGCAGCCCCTGACACGTGTTGGGGGATGTGAATGACAGCAAGCTGTTCACCGGAGCGTGGCCA  1620

Qy  1621  TTGGGGGCCCCGGAGGAGCAGGAGAACCGGGGTGCTGTTCTACCGTTCACCGGAGCCTCAG  1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  TTGGGGGCCCCGGAGGAGCAGGAGAACCGGGGTGCTGTTCTACCGTTCACCGGAGCCTCAG  1680

Qy  1681  AATCCGGCATCAGCCCCCTCCCAGCAGCGGATTGCCAGCTCTCCCCAGGC  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  AATCCGGCATCAGCCCCCTCCCAGCAGCGGATTGCCAGCTCTCCCCAGGC  1740

Qy  1741  TGCAGTATTTTGGGCAGGCGCTGAGTGGGGTCAGGGGACTGATGG  1800
          |||||||||||||||||||||||||||||||||||||||||||||
Db  1741  TGCAGTATTTTGGGCAGGCGCTGAGTGGGGTCAGGGGACTGATGG  1800

Qy  1801  ACCTGGCCGTGGGGGCCCAGGTGCTCCTGCTCAGGAGTCTGCCGGTTCTGCCGGTGCTGAAAG  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  ACCTGGCCGTGGGGGCCCAGGTGCTCCTGCTCAGGAGTCTGCCGGTTCTGCCGGTGCTGAAAG  1860

Qy  1861  TGGGGGTGGCCATGAGATTCAGGCCTGGAGGTGGCCAAGGCGTGTGTCTCCACCGTGTGTGGG  1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  TGGGGGTGGCCATGAGATTCAGGCCTGGAGGTGGCCAAGGCGTGTGTCTCCACCGTGTGTGGG  1920

Qy  1921  AAGAGAGCCCAGTGCCCTAGGTGACATCCAAAGTCTCGTCAGTTTGATCTGCACTGGACC  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  AAGAGAGCCCAGTGCCCTAGGTGACATCCAAAGTCTCGTCAGTTTGATCTGGCACTGGACC  1980

Qy  1981  GCTCACTGGACCAGCTAGGTGACAGCATCCAAAGTCTGTCGCCATTTCAATGAAACCCACTTTGACTCGAA  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  GCTCACTGGACCAGCTAGGTGACAGCATCCAAAGTCTGTCGCCATTTCAATGAAACCCACTTTGACTCGAA  2040

Qy  2041  CAGTCGTCTGACTTCTCGTCGGGGATTCACTGTGAAACCCTGAAGCTGCTTTTGCCAGATTGTG  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  CAGTCGTCTGACTTCTCGTCGGGGATTCACTGTGAAACCCTGAAGCTGCTTTTGCCAGATTGTG  2100

Qy  2101  GAAAAACCCGGGACTGGGATTCACTGTGAAACCCTGAAGCTGCTTTTGCCAGATTGTG  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  GAAAAACCCGGGACTGGGATTCACTGTGAAACCCTGAAGCTGCTTTTGCCAGATTGTG  2160

Qy  2161  TGGAGGATGGTGGTGAGCCCATCAATTCCTGCACCTCAACTTCTGT  2220
          |||||||||||||||||||||||||||||||||||||||||||||||
```

```
              GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:   June  7, 2004, 09:36:53 ; Search time 27 Seconds
                                    (without alignments)
                                    4136.236 Million cell updates/sec

Title:      US-09-891-943-2
Perfect score:  5987
Sequence:   1 TFGTVLLSVLASYHGFNLD........DTATFSGDDFSCVAPNVPLS 1161
Scoring table:  BLOSUM62
            Gapop 10.0 , Gapext 0.5

Searched:   283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters:    283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  PIR 78:*
            1: Pir1:*
            2: Pir2:*
            3: Pir3:*
            4: Pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	3818	63.8	1163	1	RWHU1C	cell surface glyco
2	3455	57.7	1153	1	RWHU1B	cell surface glyco
3	3327.5	55.6	1153	2	S00551	leukocyte surface
4	1474	24.6	1170	2	S03308	cell surface glyco
5	1457.5	24.3	1163	2	I56126	lymphocyte fuction
6	1148	19.2	1179	2	A53213	integrin alpha-E c
7	1030.5	17.2	1151	2	A45226	integrin alpha-1 c
8	1030.5	17.2	1180	2	A35854	integrin alpha-1 c
9	1012	16.9	1178	2	S44142	VLA-2 protein homo
10	1007.5	16.8	1170	2	I45914	integrin alpha-2 s
11	1007.5	16.8	1181	2	A33998	integrin alpha-2 c
12	654.5	10.9	1038	2	S06046	integrin alpha-4 c
13	646.5	10.8	1039	2	A41131	lymphocyte-Peyer's
14	638	10.7	1054	2	JC7294	alphaP integrin -
15	611	10.2	1041	2	T31437	integrin alpha cha
16	584.5	9.8	1035	2	I58409	integrin alpha-9 c
17	579.5	9.7	1053	2	I55534	VLA-3 alpha subuni
18	576.5	9.6	1051	2	A35761	cell surface glyco
19	555	9.3	1037	2	A60163	glycoprotein IIb -
20	552	9.2	1049	2	A27079	fibronectin recept
21	548	9.2	1039	2	A34269	integrin alpha-2b
22	547.5	9.1	1051	2	A40021	integrin alpha-3 c
23	535.5	8.9	1053	2	S44250	integrin alpha-5 c
24	534	8.9	1034	2	A36108	integrin alpha-V c
25	524.5	8.8	1394	2	A29637	position-specific
26	497	8.3	1044	2	T10050	integrin alpha-v c
27	496.5	8.3	1072	2	A38457	integrin alpha-6 c
28	493	8.2	1048	2	A27421	integrin alpha-5 c
29	492	8.2	1146	2	S40311	integrin - fruit f
30	477.5	8.0	1044	2	S16516	integrin alpha-8 c
31	473	7.9	1073	2	B36429	integrin alpha-6 c
32	464	7.8	1137	2	JC5950	integrin alpha-7 c
33	462.5	7.7	1091	2	A41543	integrin alpha-6 c
34	446.5	7.5	1045	2	S60571	integrin alpha v c
35	439	7.3	1135	2	I61186	alpha-7 integrin -
36	433.5	7.2	1115	2	T09433	integrin alpha cha
37	431.5	7.2	1115	2	T09403	integrin alpha cha
38	431.5	7.2	1226	2	S44824	F54F2.1 protein -
39	420.5	7.0	1106	2	S38783	integrin alpha cha
40	410.5	6.9	764	2	I36916	glycoprotein IIb -
41	394	6.6	1139	2	S28277	hypothetical prote
42	322.5	5.4	272	2	A55348	integrin alpha-1
43	319.5	5.3	126	2	B30892	leukocyte adhesion
44	319.5	5.3	1086	2	T18523	integrin alpha cha
45	313.5	5.2	604	2	I36917	glycoprotein IIb -

ALIGNMENTS

```
RESULT 1
RWHU1C
cell surface glycoprotein CD11c precursor - human
N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 22-Jun-1999
C;Accession: A36584; A35543; S00864
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A;Reference number: A36584
A;Contents: erratum
A;Accession: A36584
A;Molecule type: DNA
A;Residues: 1-1163 <COR>
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A;Reference number: A35543; MUID:9153306; PMID:2303426
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: this sequence has been revised in reference A36584
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166645; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755,'L',757-1163 <CO3>
A;Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on mye
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo]
C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1163/Product: cell surface glycoprotein CD1c #status predicted <MAT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1108-1133/Domain: transmembrane #status predicted <TMM>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status prec

Query Match        63.8%;    Score 3818; DB 1;  Length 1163;
Best Local Similarity  65.5%;    Pred. No. 6.6e-267;
Matches  755;  Conservative 138;  Mismatches 253;  Indels  6;  Gaps  5;

Qy      1 TFGTVLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT 60
```

```
Db   1082 LPGQEAFMRAQTTVLEKYKVHNPTLIVGSSIGGLLLALITAVLYKVGFFKRQYKEMM 1141
                |::: ||||
Qy   1137 EDKPEDTATFSG 1148
                |:|
Db   1142 EEANGQIAPENG 1153
```

RESULT 2
RWHU1B
cell surface glycoprotein CD11b precursor [validated] - human
N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 31-Dec-2000
C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567
R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD1
B.
A;Reference number: A31108; MUID:88315033; PMID:2457584
A;Accession: A31108
A;Molecule type: mRNA
A;Residues: 1-1153 <COR>
A;Cross-references: GB:J03925; NID:g187284; PIDN:AAA59544.1; PID:g307148
A;Note: part of this sequence was confirmed by protein sequencing
R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988
A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mc
A;Reference number: A28915; MUID:88257215; PMID:2454931
A;Accession: A28915
A;Molecule type: mRNA
A;Residues: 1-499,501-965,'P',967-1153 <ARN>
A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:g186935; PIDN:AAA594:
A;Note: the authors translated the codon TAC for residue 1129 as Thr
A;Note: part of this sequence, including the amino end of the mature protein, was confirm
R;Shelley, C.S.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991
A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg
A;Reference number: A41600; MUID:92073318; PMID:1683702
A;Accession: A41600
A;Molecule type: DNA
A;Residues: 1-9 <SHE>
A;Cross-references: GB:M76724; NID:g180018; PIDN:AAA58410.1; PID:g553215
R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988
A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi
A;Reference number: A94193; MUID:89098893; PMID:2833753
A;Accession: A30892
A;Molecule type: mRNA
A;Residues: 917-1042 <AR2>
A;Cross-references: GB:M18044
R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989
A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-int
n during evolution.
A;Reference number: A46526; MUID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <FLE>
A;Cross-references: GB:S52227; NID:g263047; PIDN:AAB24821.1; PID:g263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature n
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
```

```
Db      4 TRAALLLFTALATSLGFNLDTEELTAFRVDSAGFGDSVVQYANSWVVVGAPQKITAANQT 63
                |:|||  |:||      |:  |||||   |||  |:|||    ||||||||||||:||||
Qy     61 GRLYDCAAATGMCCQPITPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120
                |||||||| |: ||:|    ||| ||||||||||:|||::|||||||||:|::||:|||:|
Db     64 GGLYQCGVSTGACEPIGLQVPPEAVNMSLGLSLASTTSPSQLLACGPTVHHECGRNMYLT 123
                ||||:|| |:| |||  |:||||||||||||||:|::|| |||||||| | ||| :|||
Qy    121 GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE 180
                |: ||| | |: ||:|    |||  ||:|:||||||||| :|  ||:|||||:|| || |
Db    124 GLcFLLGPT-QUTQRLPVSRQBCPRQBQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQ 182
                ||  |||:|    |:| |||:||    ||:||||||||||    |||||||||||:|:||
Qy    181 GTDTLFALMQYSNLLKIHFTTQFRRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHH 240
                |: |||  |||:|||:|||  ||  |||  |||::|   |||  |: |:||:||||:|  |
Db    183 RPSTQFSLMQFSNKFQTHFTPEEFRRTSNPLSLLASVHQLCQGFTYTATAIQNVVHRLFHA 242
                |  |   |||||:|    |    ||  ||| ||:|| |:  ||: |||  |:||:|||::|
Qy    241 KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN 300
                |:|||    |||:||||||:|||||||  ||| ||||||||||||||| ||||||||:|:
Db    243 SYGARRDATKLILIVITDGKKGGSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELN 302
                ||| || :||:||||||||:|  |||||||||  | |||||||||||:|| :|:|||:||
Qy    301 TISSAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEGTQSRASSPQHEMSQEGFSTALTM 360
                |:  |||:|| ||||||:||  ||||:|||::|||||||||||  ||:|||||||||||
Db    303 DIASKPSQEHIFKVEDFDPALKDIQNQLKEKIFAIEGTETSSSFELEMAQEGFSAVFTP 362
                ||:|||||||||||||:| || ||||| |||||:|:||| ||||  |:|||||||::|
Qy    361 DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALMWKGVQNLVL 420
                ||  ||||||| ||||| |||||||||||||||||||||||||||||||||:||||:|||
Db    363 DGPVLGAVGSFTWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQSLVL 422
                |||:||||||| ||||||||||||||||||||||||||||||||||||||||||:||||
Qy    421 GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVSDGSTDLILIGAPH 480
                |||||||||||||||||||||:|:|||||||||||||||||||| ||||||| ||||||
Db    423 GAPRYQHTGKAVIFTQVSRQWRMKAEVTGTQIGSYFGASLCSVDVDTDGSTDLVLIGAPH 482
                |||||||||||||||||||||:||||||||||||||||||||||| |||||||:|||||
Qy    481 YYEQTRGGQVSVCPLPRGQRVQWOCDAVLRGEOGHPWGRFGAALTVLGDVNEDKLIDVAI 540
                ||:|||||||||||||:||| |:|||:|| |||||||||||||||||||:|:|||:||||
Db    483 YYRQTRGGQVSVCPLPRGWR-RWCDAVLYGEQGHPWGRFGAALTVLGDVNGDKLTDVVI 541
                ||:|||||||||||||| |   |||||:|||||||||||||||||||||:|||:|||:|
Qy    541 GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD 600
                ||||::||||||:|||| |||||||||||||||: |||||||||||||||||||||||:|
Db    542 GAPGBEENRGAVYLFHGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGGQDLTQDGLVD 601
                ||||:|||||||:|||  ||||||||||||||| ||| |||||||||||||||||||:|
Qy    601 LAVGARGQVYLLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLITIQKS 660
                |||||||||:|||:||||:|:||| ||| |||:||:||  |:|  || |||| ||||:|:
Db    602 LAVGARGQVYLLLRTRPVLMVGVSMQFIPAEIPRSAPECREQVVSEQTLVQSNICLYIDKR 661
                ||||||||:|||: |||:||:||: |:| |:||   ||||:|||:|||:|||||| ||:|
Qy    661 SLDQLG--DIQSSVRFDLALDPGLRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPD 718
                ||| |    ||:|  ||:||||||  |:||| ||:|||||:|||:|||| | :| ||||||
Db    662 SKNLLGSRDLQSSVTLDLALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPS 721
                ||||||||||| |  ||||||||   |||||||:||| | ||||||||| |||||||||
Qy    719 CVEDVVSPILHLNFSLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDL 778
                |||||||||:||||||||| || | ||||||| |||::|||||||||| |: ||||||
Db    722 CVEDSVTFTTLRLNFTLVGKPLLAFRNLRPMLAADAQRYFTASLPFEKNCGADHICQDNL 781
                |||| |:| ||||:|||||||||  |||||:|| ||||||||||||||||| |  |||||
Qy    779 RSLHLTCDSAPVGSQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLLLTANVS 838
                | |||||||||||| |||:||||||||:||:|| || ||||||||||||||| || ||||
Db    782 GISFSFPGLKSLLVGSNLELNAEYVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQL 841
                || |  ||||| ||||||:|||| ||:|||||||||:| ||  |||||||||:|::||||
Qy    839 SALRLACETVPTEDEGLRSSRCSVNHPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASAS 898
                | |||||| |||||||||||||| ||||||||:||||||||||||||||||||:||||||
Db    842 SENNKASSSKATFQELEPVKYAVYTMISRQEESTKYFNPATSDEKKMKEAEHRYRVNNLS 958
                |:|:|||  ||||:|||||||||||:|| ||: |||||:|||:|:|:|||||||||||||
Qy    899 GISFSFPGLKSLLVGSNLELNAEYVMVWNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGQL 841
                || |  ||||| ||||||:|||| ||:|||||||||:| ||  |||||||||:|::||||
Db    902 SENNTPRTSKTTFQLELPVKYAVYTVVSSHEQFTKYLNFSESEBKXESHVAMHRYQVNNLG 961
                |:||  |:|||||:||||||||||||||| |||||:|| |:||:||:|| |||||:||||
Qy    959 QRDLAISINFWVPVLLNGVAVM-DVVMEAP-SQSLPCVSERKPQHSDFLTQISRSPMLD 1016
                |||| |||||||:|||||| |:   |:|  |  |||  ||::|||  |:| |:|||||||
Db    962 QRDLPVSINFWVPVELNQEAVWMDVEVSHPQNPSLRCSSEKIIAPPASDFLAHIQKNPVLD 1021
                |||| |||||||:||||:|||||||:|   |: ||||| |:|||||||:|| |||||||||
Qy   1017 CSIADCLQFRCDVPSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQ 1076
                |:|| |||||||||||:|||||||||||||||||||||||||||||||||||:||||||||
Db   1022 CSIAGCLRFRCDVPSFSVQEELDFTLKGNLSFGWVRQILQKKVSVVSVAEITFDTSVISQ 1081
                ||||||:|||||||||||||||||||||||||||||||:||||||||||||||||||:||
Qy   1077 LPGQEAFMRAQEMEMVLEEDEVVNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEML 1136
                ||||||||||||:|:|  |:|  |||||:||||:||||||||:|||||||||||:|||||:
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:      June  7, 2004, 09:33:03 ; Search time 17 Seconds
                          (without alignments)
                          3556.085 Million cell updates/sec

Title:         US-09-891-943-2
Perfect score: 5987
Sequence:      1 TFGTVLLSVLASYHGFNLD.........DTATFSGDDFSCVAPNVPLS 1161

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:      141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     SwissProt_42:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
```

```
Result          %       Query
No.   Score   Match   Length  DB  ID            Description
----- -----   -----   ------  --  ------------  -----------
 1    5987    100.0   1162    1   ITAD_HUMAN    Q13349 homo sapien
 2    3815    63.7    1163    1   ITAX_HUMAN    P20702 homo sapien
 3    3449.5  57.4    1152    1   ITAM_HUMAN    P11215 homo sapien
 4    3327.5  55.6    1153    1   ITAM_MOUSE    P05555 mus musculu
 5    1476    24.7    1170    1   ITAL_HUMAN    P20701 homo sapien
 6    1457.5  24.3    1163    1   ITAL_MOUSE    P24063 mus musculu
 7    1156.5  19.3    1167    1   ITAE_MOUSE    Q60677 mus musculu
 8    1148    19.2    1179    1   ITAE_HUMAN    P38570 homo sapien
 9    1146.5  19.1    1189    1   ITAH_HUMAN    Q9ukx5 homo sapien
10    1131.5  18.9    1167    1   ITAG_HUMAN    O75578 homo sapien
11    1030.5  17.2    1151    1   ITA1_HUMAN    P56199 homo sapien
12    1030.5  17.2    1180    1   ITA1_RAT      P18614 rattus norv
13    1012    16.9    1178    1   ITA2_MOUSE    Q62469 mus musculu
14    1007.5  16.8    1170    1   ITA2_BOVIN    P53710 bos taurus
15    1007.5  16.8    1181    1   ITA2_HUMAN    P17301 homo sapien
16    654.5   10.9    1038    1   ITA4_HUMAN    P13612 homo sapien
17    646.5   10.8    1039    1   ITA4_MOUSE    Q00651 mus musculu
18    585.5    9.8    1032    1   ITA4_XENLA    Q91687 xenopus lae
19    584.5    9.8    1035    1   ITA9_HUMAN    Q13797 homo sapien
20    579.5    9.7    1053    1   ITA3_MOUSE    Q62470 mus musculu
21    570.5    9.5    1066    1   ITA3_CRISP    P17852 cricetidae
22    554      9.3    1049    1   ITA5_HUMAN    P08648 homo sapien
23    547      9.1    1039    1   ITAB_HUMAN    P08514 homo sapien
24    540.5    9.0    1066    1   ITA3_HUMAN    P26006 homo sapien
25    538      9.0    1033    1   ITAB_MOUSE    Q9qum0 mus musculu
26    535.5    8.9    1053    1   ITAV_MOUSE    P11688 mus musculu
27    534      8.9    1034    1   ITAV_CHICK    P26008 gallus gall
28    534      8.9    1050    1   ITA5_XENLA    Q06274 xenopus lae
29    525.5    8.8    1396    1   ITA2_DROME    P12080 drosophila
30    497      8.3    1044    1   ITAV_MOUSE    P43406 mus musculu
31    496.5    8.3    1072    1   ITA6_CHICK    P26007 gallus gall
32    493      8.2    1048    1   ITAV_HUMAN    P06756 homo sapien
33    490      8.2    1146    1   ITA1_DROME    Q24247 drosophila
34    486      8.1    1130    1   ITA6_HUMAN    P23229 homo sapien
35    482.5    8.1    1025    1   ITA8_HUMAN    P53708 homo sapien
36    477.5    8.0    1044    1   ITA8_CHICK    P26009 gallus gall
37    455      7.6    1181    1   ITA7_HUMAN    Q13683 homo sapien
38    437.5    7.3    1091    1   ITA6_MOUSE    Q61739 mus musculu
39    431.5    7.2    1226    1   PAT2_CAEEL    P34446 caenorhabdi
40    430.5    7.2    1179    1   ITA_MOUSE     Q61738 mus musculu
41    429.5    7.2    1115    1   ITA3_DROME    O44386 drosophila
42    420.5    7.0    1106    1   ITA7_RAT      Q63258 rattus norv
43    411.5    6.9    1000    1   ITA5_DROME    Q9wim8 drosophila
44    408.5    6.8    1015    1   ITA4_DROME    Q9v7a4 drosophila
45    394      6.6    1139    1   INA1_CAEEL    Q03600 caenorhabdi
```

                                    ALIGNMENTS

```
RESULT 1
ITAD_HUMAN
ID   ITAD_HUMAN      STANDARD;      PRT;   1162 AA.
AC   Q13349; Q15575; Q15576;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN   ITGAD.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=96111956; PubMed=8777714;
RA   Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA   Staunton D.E., Gallatin W.M.;
RT   "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT   3.";
RL   Immunity 3:683-690(1995).
RN   [2]
RP   SEQUENCE OF 1-235 FROM N.A.
RX   MEDLINE=20187620; PubMed=10722744;
RA   Noti J.D., Johnson A.K., Dillon J.D.;
RT   "Structural and functional characterization of the leukocyte integrin
RT   gene CD11d. Essential role of Sp1 and Sp3.";
RL   J. Biol. Chem. 275:8959-8969(2000).
RN   [3]
RP   SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX   MEDLINE=96257236; PubMed=8666289;
RA   Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT   "Cloning and chromosomal localization of a novel gene-encoding a human
RT   beta 2-integrin alpha subunit.";
RL   Gene 171:291-294(1996).
RN   [4]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99059842; PubMed=9841932;
RA   Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,
RA   Hoffman P.A., Staunton D.E., Bochner B.S.;
RT   "alphadbeta2 integrin is expressed on human eosinophils and functions
RT   as an alternative ligand for vascular cell adhesion molecule 1
RT   (VCAM-1).";
RL   J. Exp. Med. 188:2187-2191(1998).
RN   [5]
RP   INTERACTION WITH VCAM1.
RX   MEDLINE=99370002; PubMed=10438935;
RA   Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA   Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT   "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
RT   binding interface between I domain and VCAM-1.";
RL   J. Immunol. 163:1984-1990(1999).
CC   -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND
CC       VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS
CC       CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-
```

```
CC         BORNE PATHOGENS, PARTICULATE MATTER, AND SENESCENT ERYTHROCYTES
CC         FROM THE BLOOD.
CC      -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
CC         ASSOCIATES WITH BETA-2.
CC      -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -!- TISSUE SPECIFICITY: EXPRESSED MODERATELY ON MYELOMONOCYTIC CELL
CC         LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
CC         TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN
CC         ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
CC      -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC         WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC      -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC      -!- SIMILARITY: Contains 1 VWFA domain.
CC      -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC   ---------------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute.  There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U37028; AAB38547.1; -.
DR   EMBL; U40274; AAB60634.1; -.
DR   EMBL; U40275; AAB60635.1; -.
DR   EMBL; U40276; AAB60636.1; -.
DR   EMBL; U40277; AAB60637.1; -.
DR   EMBL; U40279; AAB60638.1; -.
DR   EMBL; U40278; AAB60638.1; JOINED.
DR   EMBL; AF187881; AAF62875.1; -.
DR   HSSP; P11215; 1A8X.
DR   Genew; HGNC:6146; ITGAD.
DR   MIM; 602453; -.
DR   GO; GO:0008305; C:integrin complex; TAS.
DR   GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR   GO; GO:0016337; P:cell-cell adhesion; NAS.
DR   GO; GO:0007160; P:cell-matrix adhesion; NAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; integrin_A; 1.
DR   Pfam; PF00092; vwa; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 4.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW   Signal; Repeat; Calcium;
KW   Magnesium.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18   1162       INTEGRIN ALPHA-D.
FT   DOMAIN       18   1100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1101   1124       POTENTIAL.
FT   DOMAIN     1125   1162       CYTOPLASMIC (POTENTIAL).
FT   REPEAT       32     85       FG-GAP 1.
FT   REPEAT        ?      ?       FG-GAP 2.
FT   DOMAIN      150    332       VWFA.
FT   REPEAT      350    400       FG-GAP 3.
FT   REPEAT      401    452       FG-GAP 4.
FT   REPEAT      454    516       FG-GAP 5.
FT   REPEAT      518    576       FG-GAP 6.
FT   REPEAT      581    633       FG-GAP 7.
FT   CA_BIND     465    473       POTENTIAL.
FT   CA_BIND     530    538       POTENTIAL.
FT   CA_BIND     593    601       POTENTIAL.
FT   SITE       1127   1131       GFFKR MOTIF.
FT   DISULFID     67     74       BY SIMILARITY.
FT   DISULFID    106    124       BY SIMILARITY.
FT   DISULFID    655    710       BY SIMILARITY.
FT   DISULFID    769    775       BY SIMILARITY.
FT   DISULFID    846    861       BY SIMILARITY.
FT   DISULFID    994   1018       BY SIMILARITY.
FT   DISULFID   1023   1028       BY SIMILARITY.
FT   CARBOHYD     59     59       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD     87     87       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD     99     99       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD    391    391       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD    691    691       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD    733    733       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD    873    873       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD    957    957       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CARBOHYD   1046   1046       N-LINKED (GLCNAC...)  (POTENTIAL).
FT   CONFLICT    500    500       MISSING (IN REF. 2).
FT   CONFLICT    515    518       GHPW -> ATP (IN REF. 2).
FT   CONFLICT    825    825       L -> V (IN REF. 2).
FT   CONFLICT    984    984       V -> A (IN REF. 2).
SQ   SEQUENCE   1162 AA;  126885 MW;  F296A1A35455D77D CRC64;
```

```
Query Match            100.0%;   Score 5987;   DB 1;  Length 1162;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 1161;        Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1   TPFGTVLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2   TPFGTVLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT   61

Qy    61   GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    62   GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK  121

Qy   121   GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIQNDFNQMKGFVQAVMGQFE  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122   GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIQNDFNQMKGFVQAVMGQFE  181

Qy   181   GTDTLFALMQYSNLLKIHFFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVTQLFHH  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182   GTDTLFALMQYSNLLKIHFFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVTQLFHH  241

Qy   241   KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242   KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN  301

Qy   301   TISSAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302   TISSAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  361

Qy   361   DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   362   DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL  421

Qy   421   GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLIIGAPH  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   422   GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLIIGAPH  481

Qy   481   YYEQTRGGQVSVCPLRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   482   YYEQTRGGQVSVCPLRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI  541

Qy   541   GAPGEQENRGAVLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   542   GAPGEQENRGAVLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD  601

Qy   601   LAVGARGQVLLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLLTIQKS  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   602   LAVGARGQVLLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLLTIQKS  661

Qy   661   SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPLTRRKTLGLIHCETLKLLPDCV  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   662   SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPLTRRKTLGLIHCETLKLLPDCV  721

Qy   721   EDVVSPIILHLNFSLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGV  780
           |||||||||||||||||||||||||||||||||||
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on:       June  7, 2004, 09:36:23 ; Search time 56 Seconds
                                         (without alignments)
                                         6541.368 Million cell updates/sec

Title:        US-09-891-943-2
Perfect score: 5987
Sequence:     1 TFGTVLLLSVLASYHGFNLD.........DTATFSGDDFSCVAPNVPLS 1161

Scoring table: BLOSUM62
              Gapop 10.0 , Gapext 0.5

Searched:     1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:    1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    SPTREMBL 25:*
              1: sp_archea:*
              2: sp_bacteria:*
              3: sp_fungi:*
              4: sp_human:*
              5: sp_invertebrate:*
              6: sp_mammal:*
              7: sp_mhc:*
              8: sp_organelle:*
              9: sp_phage:*
             10: sp_plant:*
             11: sp_rodent:*
             12: sp_virus:*
             13: sp_vertebrate:*
             14: sp_unclassified:*
             15: sp_rvirus:*
             16: sp_bacteriap:*
             17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB	ID	Description
1	4415.5	73.8	1161	11	Q9QYE7	Q9qye7 rattus norv
2	3824	63.9	1169	4	Q8IVA6	Q8iva6 homo sapien
3	3581.5	59.8	1169	11	Q9QXH4	Q9qxh4 mus musculu
4	3314	55.4	1151	11	Q9JI30	Q9ji30 rattus norv
5	2829	47.3	920	6	Q28984	Q28984 sus scrofa
6	2755	46.0	1036	11	Q8CA73	Q8ca73 mus musculu
7	1451.5	24.2	1161	11	Q9WTV4	Q9wtv4 mus musculu
8	1447	24.2	1160	11	Q9R200	Q9r200 mus musculu
9	1318	22.0	1086	4	Q96H81	Q96hb1 homo sapien
10	1297.5	21.7	1196	13	Q98TF1	Q98tf1 cyprinus ca
11	1263	21.1	1187	13	Q98TF0	Q98tf0 cyprinus ca
12	1220.5	20.4	927	6	Q8HZV0	Q8hzv0 bos taurus
13	1164.5	19.5	1167	11	O88340	O88340 rattus norv
14	1138.5	19.0	1167	11	O88341	O88341 rattus norv
15	1096.5	18.3	1188	11	Q7TQC3	Q7tqc3 mus musculu
16	1041	17.4	1171	13	O42094	O42094 gallus gall
17		17.0	1017	4	Q8TES5	Q8tes5 homo sapien
18		17.0	1015.5	1038	11	Q8BS01	Q8bs01 mus musculu
19		16.4	984.5	895	11	Q9WUF8	Q9wuf8 mus sp. itg
20		16.2	972.5	1160	6	Q8MKF4	Q8mkf4 felis silve
21		13.6	812	4	Q9BPQ8	Q9bpq8 halocynthia
22		12.9	771.5	823	4	Q8WYI8	Q8wyi8 homo sapien
23		12.1	722	11	Q8CE84	Q8ce84 mus musculu
24		11.0	656.5	1032	11	Q61989	Q61989 mus musculu
25		10.7	643	13	Q06271	Q06271 xenopus lae
26		10.7	638	5	Q9U6S1	Q9u6s1 strongyloce
27		10.7	629.5	1474	5	Q86G87	Q86g87 pseudoplusi
28		10.5	629	11	O63001	O63001 rattus norv
29		10.5	622	6	Q9BGU3	Q9bgu3 bos taurus
30		10.4	621.5	11	Q91YD5	Q91yd5 mus musculu
31		10.4	620	13	Q98TT7	Q98tt7 gallus gall
32		10.3	615	5	Q9UB90	Q9ub90 lytechinus
33		10.2	611	5	O76378	O76378 lytechinus
34		10.2	610	11	Q8C270	Q8c270 mus musculu
35		9.5	568.5	1036	5	Q7YRP8	Q7yrp8 equus caba1
36		9.5	548	1034	6	Q9TUN4	Q9tun4 oryctolagus
37		9.2	537	1016	13	Q91779	Q91779 xenopus lae
38		9.0	532.5	1053	11	Q8OYP5	Q8oyp5 mus musculu
39		8.8	526	1119	5	Q8GG88	Q8gg88 pseudoplusi
40		8.5	509	231	4	Q8N882	Q8n882 homo sapien
41		8.4	502.5	1012	11	O70304	O70304 mus musculu
42		8.2	490	1007	6	Q9GK48	Q9gk48 bos taurus
43		8.2	490	1033	13	O42598	O42598 xenopus lae
44		8.2	490	1049	5	Q8SY51	Q8sy51 drosophila
45		8.2	488	1047	6	Q9MZD6	Q9mzd6 bos taurus

ALIGNMENTS

```
RESULT 1
Q9QYE7
ID   Q9QYE7         PRELIMINARY;     PRT;   1161 AA.
AC   Q9QYE7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alpha D integrin.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA   Gallatin W.M.;
RT   "Cloning of rat alpha D, a novel beta 2 integrin.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF021334; AAF21241.1; -.
DR   HSSP; P11215; 1BHO.
DR   GO; GO:0008305; C:integrin complex; IEA.
DR   GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR   GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR   GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; integrin_A; 1.
DR   Pfam; PF00092; vwa; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00191; Int_alpha; 4.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Integrin.
SQ   SEQUENCE   1161 AA;  126600 MW;  2258491A98A70SE CRC64;
```

```
Db  1021  DCLHFRCDIPSLDIQDELDFILRGNLSFGWVSQTLQEKVLLVSEABITFDTSVYSQLPGQ  1080
          :|| :|||||||  ||||  ||:| :||||||:|  |:|||||: |:  |||||||||||
Qy  1081  EAFMRAQMEMVLEEDEVNAIPIIMGGSVGALLLALITATLYKLGFFKRHYKEMLEDKP   1140
          ||| :|||  :||| :::  | |  |:||| |||||||:: ||||||||:||||:|  :
Db  1081  EAFLRAQVETTLEEYVVYEPIFLVAGSSVGGLLLLALITVVLYKLGFFKRQYKEMLDGKA  1140

Qy  1141  EDTATFSGDDFSCVAP  1156
          :| :|  | ||| | |
Db  1141  ADPVTAGQADFGCETP  1156
```

RESULT 2
Q8IVA6
```
ID   Q8IVA6        PRELIMINARY;      PRT;   1169 AA.
AC   Q8IVA6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Integrin, alpha X (Antigen CD11C (p150), alpha polypeptide).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RA   Strausberg R.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC038237; AAH38237.1; -.
DR   GO; GO:0008305; C:integrin complex; IEA.
DR   GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR   GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR   InterPro; IPR000413; Integrin_alpha.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF01839; FG-GAP; 3.
DR   Pfam; PF00357; integrin_A; 1.
DR   Pfam; PF00092; vwa; 1.
DR   PRINTS; PR01185; INTEGRINA.
DR   PRINTS; PR00453; VWFADOMAIN.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR   PROSITE; PS50234; VWFA; 1.
SQ   SEQUENCE 1169 AA; 128521 MW; A17B484FEFC79EB6 CRC64;
```

```
Query Match             63.9%;  Score 3824;  DB 4;  Length 1169;
Best Local Similarity   65.8%;  Pred. No. 4.6e-285; -
Matches  758;  Conservative 135;  Mismatches 253;  Indels 6;  Gaps 5;
```

```
Qy    1  TFGTVLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT   60
Db    4  TRAALLFTALATSLGFNLDTEELTAFRVDSAGFGDSVVQYANSWVVGAPQKITAANQT    63

Qy   61  GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK  120
Db   64  GGLYQCGYSTGACEPIGLQVPPEAVNMSLGLSLASTSPSQLLACGPTVHHECGRNMYLT   123

Qy  121  GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE  180
Db  124  GLCFLLGPT-QLTQRLPVSRQECPRQQDIVFLIDGSGSISSRNFATMMNFVRAVISQFQ   182

Qy  181  GTDTLFALMQYSNLLKIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHH  240
Db  183  RPSTQFSLMQFSNKFQTHFTFEEFRRSNPLSILASVHQLQGFTYTAIQNVHRLFHA     242

Qy  241  KNGARKSAKKILIVITDGQKYDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN   300
Db  243  SYGARRDAKILIVITDGKKEGSLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELN    302

Qy  301  TISSAPPQDHVFKVDNPAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  360
Db  303  DIASKPSQEHIFKVEDFPDALKDIQNQLKEKIFAIEGTETTSSSSFELEMAQEGFSAVFTP 362
```

```
Query Match             73.8%;  Score 4415.5;  DB 11;  Length 1161;
Best Local Similarity   73.8%;  Pred. No. 0; -
Matches  853;  Conservative 111;  Mismatches 187;  Indels 5;  Gaps 4;
```

```
Qy    3  GTVLLS--VLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT   60
Db    4  GVIILLCGWVIASCHGSNLDVEERIVFRRDAASFGQTVVQFGGSRLVVGAPLEAVAVNQT   63

Qy   61  GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK  120
Db   64  GRLYDCAPATGMCQPIVLRSPLEAVNMSLGLSLVTATNNAQLLACGPTAQRACVKNMTAK  123

Qy  121  GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE  180
Db  124  GSCLLLGSSLQFIQAVPASMPECPRQEMDIAFLIDGSGSINQRDFAQMKDFVKALMGEFA  183

Qy  181  GTDTLFALMQYSNLLKIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHH  240
Db  184  STSTLPSLMQYSNLIKTHFTFTEFKNILDPQSLVDPIVQLQGLTYTATGIRTVMEELFHS  243

Qy  241  KNGARKSAKKILIVITDGQKYDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN   300
Db  244  KNGSRKSAKKILIVTTDGQKYRDPLEYSDVIPAADKAGIIRYAIGVGDAFQEPTALKELN  303

Qy  301  TISSAPPQDHVFKVDNPAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  360
Db  304  TIGSAPPQDHVFKVGNFAALRSIQRQLQEKIFAIEGTQSRSSSSFQHEMSQEGFSSALTS  363

Qy  361  DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL  420
Db  364  DGPVLGAVGSFSWSGGAFLYPPNTRPTFINMSQENVDMRDSYLGYSTAVAFWKCVHSLIL  423

Qy  421  GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPH  480
Db  424  GAPRHQHTGKVVIFTQEARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSTDLVLIGAPH  483

Qy  481  YYEQTRGGQVSVCPLPRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI  540
Db  484  YYEQTRGGQVSVFVPVP-GVRGRWQCEATLHGEQGHPWGRFGVALTVLGDVNGDNLADVAI 542

Qy  541  GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD  600
Db  543  GAPGEESRGAVYIFHGASRLEIMSPSQRVTGSQLSRLQYFGQSLSGGQDLTQDGLVD    602

Qy  601  LAVGARGQVLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKS   660
Db  603  LAVGAQGHVLLLRSLPLLKVELSIRFAPMEVAKAVYQCWERTPTVLEAGEATVCLTVHKG  662

Qy  661  SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPTLTRRKTLGGLIHCETLKLLLPDCV  720
Db  663  SPDLLGNVQGSVRYDLALDPGRLISRAIFDETKNCTLTGRKTLGLGDHCETVKLLLPDCV  722

Qy  721  EDVVSPIIHLNFSLVRFPIPSPQNLRPVLAVGSQDLFTASLPFKNCGQDGLCEGDLGV   780
Db  723  EDAVSPIILRLNFSLVRDS-ASPRNLHPVLAVGSQDHITASLPFEKNCKQELLCEGDLGI  781

Qy  781  TLSFSGLQTLTVGSSLELNVITVTWNAGEDSYGTVVSLLYYPAGLSHRRVSGAQKQPHQSA 840
Db  782  SFNFSGLQVLVVGGSPELTVTVTWNEGEDSYGTLVKFYYPAGLSYRRVTGTQ-QPHQYP  840

Qy  841  LRLACETVPTEBDEGLRSSRCSVNHPIFHEGSNGTPIVTFPVSYKATLGDRMLMRASASSE 900
Db  841  LRLACEAEPAAQEDLRSSSCSINHPIFREGAKTFMITFPVSYKAFLGDRLLLRAKASSE  900

Qy  901  NNKASSSKATFQLELPVKYAVYTMLSRQEESTKYPNFATSDEKKMKEAHRYRVNNLSQR   960
Db  901  NNKPDTNKTAFQLELPVKYTVYTLISRQEDSTNHVNFSSSHGGRQEAAHRYRVNNLSPL   960

Qy  961  DLAISINFWFVPLLNGVAVNDVVMEAPSQSLPCVSERKPPQHSDFLTQISRSPMLDCSIA 1020
Db  961  KLAVRVNFWFVPLLNGVAVNDVTLSSPAQGVSCVSQMKPPQNPDFLTQIQRRSVLDCSIA 1020

Qy 1021  DCLQFRCDVPSFSVQEBLDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQ 1080
```

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:    June  7, 2004, 09:32:28 ; Search time 65 Seconds
                                     (without alignments)
                                     5046.733 Million cell updates/sec

Title:          US-09-891-943-2
Perfect score:  5987
Sequence:       1 TFGTVLLSVLASYHGFNLD.........DTATFSGDDFSCVAPNVPLS 1161

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:          1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_29Jan04:*
             1: geneseqp1980s:*
             2: geneseqp1990s:*
             3: geneseqp20008:*
             4: geneseqp2001s:*
             5: geneseqp2002s:*
             6: geneseqp2003as:*
             7: geneseqp2003bs:*
             8: geneseqp2004s:*

             Pred. No. is the number of results predicted by chance to have a
             score greater than or equal to the score of the result being printed,
             and is derived by analysis of the total score distribution.

                                    SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5987	100.0	1161	2	AAR78166	Aar78166 Human bet
2	5987	100.0	1161	2	AAW23049	Aaw23049 Human bet
3	5987	100.0	1161	2	AAW57491	Aaw57491 Human bet
4	5987	100.0	1161	2	AAW72825	Aaw72825 Human alp
5	5987	100.0	1161	3	AAB07359	Aab07359 Human alp
6	5987	100.0	1161	5	ABG61468	Abg61468 Human Bet
7	5981	99.9	1161	2	AAW65089	Aaw65089 Human Bet
8	5981	99.9	1161	2	AAW73342	Aaw73342 Human alp
9	5971.5	99.7	1161	2	AAW23064	Aaw23064 Human bet
10	5971.5	99.7	1161	2	AAW65106	Aaw65106 Human Bet
11	5971.5	99.7	1161	2	AAW72837	Aaw72837 Human alp
12	5971.5	99.7	1161	2	AAW73343	Aaw73343 Human alp
13	5971.5	99.7	1161	3	AAB07376	Aab07376 Human alp
14	5971.5	99.7	1161	5	ABG61485	Abg61485 Human Bet
15	4403.5	73.6	1161	2	AAW23062	Aaw23062 Rat beta
16	4403.5	73.6	1161	2	AAW60004	Aaw60004 Rat alpha
17	4403.5	73.6	1161	2	AAW72824	Aaw72824 Rat alpha
18	4403.5	73.6	1161	3	AAB07374	Aab07374 Rat alpha
19	4403.5	73.6	1161	5	ABG61483	Abg61483 Rat Beta2
20	4397.5	73.5	1161	2	AAW65104	Aaw65104 Rat beta-
21	4397.5	73.5	1161	2	AAW73345	Aaw73345 Rat alpha
22	4395.5	73.4	1161	2	AAR78169	Aar78169 Rat alpha
23	4365.5	72.9	1151	2	AAR78179	Aar78179 Rat beta
24	4364.5	72.9	1151	2	AAW23059	Aaw23059 Rat beta
25	4364.5	72.9	1151	2	AAW60001	Aaw60001 Rat alpha
26	4364.5	72.9	1151	2	AAW65101	Aaw65101 Rat beta-
27	4364.5	72.9	1151	2	AAW72834	Aaw72834 Rat alpha
28	4364.5	72.9	1151	2	AAW73344	Aaw73344 Rat alpha
29	4364.5	72.9	1151	3	AAB07371	Aab07371 Rat alpha
30	4364.5	72.9	1151	5	ABG61480	Abg61480 Rat Beta2
31	4305.5	71.9	1161	2	AAW23061	Aaw23061 Mouse bet
32	4305.5	71.9	1161	2	AAW60003	Aaw60003 Mouse alp
33	4305.5	71.9	1161	2	AAW65103	Aaw65103 Mouse bet
34	4305.5	71.9	1161	2	AAW72836	Aaw72836 Mouse bet
35	4305.5	71.9	1161	2	AAW73347	Aaw73347 Mouse alp
36	4305.5	71.9	1161	3	AAB07373	Aab07373 Mouse alp
37	4305.5	71.9	1161	5	ABG61482	Abg61482 Mouse Bet
38	4303.5	71.9	1161	2	AAR78168	Aar78168 Mouse alp
39	4278.5	71.5	1155	2	AAR78167	Aar78167 Mouse alp
40	4278.5	71.5	1155	2	AAW23060	Aaw23060 Mouse bet
41	4278.5	71.5	1155	2	AAW60002	Aaw60002 Mouse alp
42	4278.5	71.5	1155	2	AAW65102	Aaw65102 Mouse bet
43	4278.5	71.5	1155	2	AAW72835	Aaw72835 Mouse alp
44	4278.5	71.5	1155	2	AAW73346	Aaw73346 Mouse alp
45	4278.5	71.5	1155	3	AAB07372	Aab07372 Mouse alp

```
                                   ALIGNMENTS

RESULT 1
AAR78166
ID   AAR78166 standard; protein; 1161 AA.
XX
AC   AAR78166;
XX
DT   28-DEC-1995  (first entry)
XX
DE   Human beta-2 integrin alpha-d.
XX
KW   Beta-2 integrin alpha-d subunit; antiinflammatory; arteriosclerosis;
KW   inflammatory bowel disease; asthma.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          17..1108
FT                   /note= "extracellular domain"
FT   Region          150..352
FT                   /note= "this region is homologous to the insert common to
FT                   CHl1a,b,c and may be a site for interaction with ICAM
FT                   family proteins"
FT   Binding-site    465..474
FT                   /note= "putative cation binding site"
FT   Binding-site    518..527
FT                   /note= "putative cation binding site"
FT   Binding-site    592..600
FT                   /note= "putative cation binding site"
FT   Region          1109..1128
FT                   /note= "transmembrane region"
FT   Domain          1129..1161
FT                   /note= "cytoplasmic domain"
XX
PN   WO9517412-A1.
XX
PD   29-JUN-1995.
XX
PF   21-DEC-1994;  94WO-US014832.
XX
PR   23-DEC-1993;  93US-00173497.
PR   05-AUG-1994;  94US-00286889.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM,  Van Der Vieren M;
XX
DR   WPI; 1995-240603/31.
DR   N-PSDB; AAQ91712.
```

```
XX
PT   Alpha sub-unit polypeptide of human beta 2 integrin - used to identify
PT   potential antiinflammatory agents, for the treatment of graft
PT   arteriosclerosis, inflammatory bowel disease, asthma, etc.
XX
PS   Claim 7; Page 82-87; 172pp; English.
XX
CC   A probe based on a partial cDNA clone (given in AAQ91727) of canine alpha
CC   -TM1 was used to screen a human spleen cDNA library to identify clone
CC   19A2 encoding the beta-1 integrin alpha-subunit, alpha-d. The cDNA was
CC   manipulated to allow expression of recombinant alpha-d subunit in COS and
CC   CHO cells
XX
SQ   Sequence 1161 AA;

Query Match            100.0%;   Score 5987;   DB 2;   Length 1161;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 1161;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1  TFGTVLLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  TFGTVLLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT   60

Qy   61  GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK  120

Qy  121  GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE  180

Qy  181  GTDTLFALMQYSNLLKIHFTFTQPRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHH  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GTDTLFALMQYSNLLKIHFTFTQPRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHH  240

Qy  241  KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN  300

Qy  301  TISSAPPQDHVFKVDNFAALGSIQKQLQBKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  TISSAPPQDHVFKVDNFAALGSIQKQLQBKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  360

Qy  361  DGLFLGAVGSFSWSGGAPLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  DGLFLGAVGSFSWSGGAPLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL  420

Qy  421  GAPRYQHTGKAVITTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPH  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GAPRYQHTGKAVITTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPH  480

Qy  481  YYEQTRGGQVSVCPLPRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  YYEQTRGGQVSVCPLPRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI  540

Qy  541  GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD  600

Qy  601  LAVGARGQVLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKS  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  LAVGARGQVLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKS  660

Qy  661  SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLPDCV  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLPDCV  720

Qy  721  EDVVSPIILHLNFSLVREPIPSPQNLRPVLAVGSQDLFTASLPPEKNCGQDGLCEGDLGV  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  EDVVSPIILHLNFSLVREPIPSPQNLRPVLAVGSQDLFTASLPPEKNCGQDGLCEGDLGV  780

Qy  781  TLSFSGLQTLTVGSSLELNVIVTWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSA  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  TLSFSGLQTLTVGSSLELNVIVTWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSA  840

Qy  841  LRLACETVPTEDEGLRSSRCSVNHPIFHEGSNGTFIVTPDVSYKATLGDRMLMRASASSE  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  LRLACETVPTEDEGLRSSRCSVNHPIFHEGSNGTFIVTPDVSYKATLGDRMLMRASASSE  900

Qy  901  NNKASSSKATFQLELPVKYAYYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQR  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  NNKASSSKATFQLELPVKYAYYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQR  960

Qy  961  DLAISINFWVPVLLNGVAVWDVVMEAPSQSLPCVSERKPPQHSDFLTQISRSPMLDCSIA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  DLAISINFWVPVLLNGVAVWDVVMEAPSQSLPCVSERKPPQHSDFLTQISRSPMLDCSIA 1020

Qy 1021  DCLQFRCDVPSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTVSYSQLPGQ 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  DCLQFRCDVPSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTVSYSQLPGQ 1080

Qy 1081  EAFMRAQMEMVLEEDEVYNAIPIIMGSSVGALLLALITATLYKLGFFKRHYKEMLEDKP  1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  EAFMRAQMEMVLEEDEVYNAIPIIMGSSVGALLLALITATLYKLGFFKRHYKEMLEDKP  1140

Qy 1141  EDTATFSGDDFSCVAPNVPLS  1161
         |||||||||||||||||||||
Db 1141  EDTATFSGDDFSCVAPNVPLS  1161


RESULT 2
AAW23049
ID   AAW23049 standard; protein; 1161 AA.
XX
AC   AAW23049;
XX
DT   24-FEB-1998  (first entry)
XX
DE   Human beta 2 integrin alpha d subunit.
XX
KW   Beta 2 integrin alpha d subunit; human; cell migration; cell adhesion;
KW   phagocytosis; diabetes; atherosclerosis; multiple sclerosis; asthma;
KW   psoriasis; lung inflammation; acute respiratory distress syndrome;
KW   rheumatoid arthritis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          17..1108
FT                   /label= Extracellular_domain
FT   Domain          150..352
FT                   /note= "region homologous to the I (insertion) domain
FT                   common to CD11a, CD11b and CD11c"
FT   Domain          1109..1128
FT                   /label= Transmembrane_domain
FT                   /note= "homologous to the human CD11c transmembrane
FT                   region"
FT   Domain          1129..1161
FT                   /label= Cytoplasmic_domain
XX
PN   WO9731099-A1.
XX
PD   28-AUG-1997.
XX
PF   24-FEB-1997;  97WO-US002713.
XX
PR   22-FEB-1996;  96US-00605672.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Gallatin WM, Van Der Vieren M;
XX
DR   WPI; 1997-435154/40.
DR   N-PSDB; AAT79220.
XX
PT   Hybridoma 199M and antibody secreted by it - specific for new rat beta2
```

OM protein - protein search, using sw model

```
            GenCore version 5.1.6
    Copyright (c) 1993 - 2004 Compugen Ltd.
```

Run on: June 7, 2004, 09:40:59 ; Search time 54 Seconds
                            (without alignments)
                6048.782 Million cell updates/sec

Title: US-09-891-943-2
Perfect score: 5987
Sequence: 1 TFGTVLLSVLASYHGFNLD..........DTATFSGDDFSCVAPNVPLS 1161

Scoring table: BLOSUM62
            Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : Published_Applications_AA:*
```
 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5987	100.0	1161	9	US-09-350-259-2	Sequence 2, Appli
2	5987	100.0	1161	10	US-09-891-943-2	Sequence 2, Appli
3	5971.5	99.7	1161	9	US-09-350-259-99	Sequence 99, Appl
4	5971.5	99.7	1161	10	US-09-891-943-99	Sequence 99, Appl
5	4403.5	73.6	1161	10	US-09-350-259-55	Sequence 55, Appl
6	4403.5	73.6	1161	9	US-09-350-259-55	Sequence 55, Appl
7	4364.5	72.9	1151	10	US-09-891-943-37	Sequence 37, Appl
8	4364.5	72.9	1151	9	US-09-350-259-37	Sequence 37, Appl
9	4305.5	71.9	1161	9	US-09-350-259-53	Sequence 53, Appl
10	4305.5	71.9	1161	9	US-09-891-943-53	Sequence 53, Appl
11	4278.5	71.5	1155	9	US-09-350-259-46	Sequence 46, Appl
12	4278.5	71.5	1155	10	US-09-891-943-46	Sequence 46, App
13	3815	63.7	1163	15	US-10-116-275-204	Sequence 204, App
14	3746	62.6	1163	9	US-09-350-259-4	Sequence 4, Appli
15	3746	62.6	1163	10	US-09-891-943-4	Sequence 4, Appli
16	3455	57.7	1153	9	US-09-350-259-3	Sequence 3, Appli
17	3455	57.7	1153	10	US-09-902-481A-1	Sequence 1, Appli
18	3455	57.7	1153	10	US-09-891-943-3	Sequence 3, Appli
19	3455	57.7	1153	14	US-10-144-259-30	Sequence 30, Appl
20	3455	57.7	1153	14	US-10-207-655-176	Sequence 176, App
21	3439.5	57.4	1152	9	US-09-945-265-4	Sequence 4, Appli
22	3417	57.1	1137	10	US-09-902-481A-6	Sequence 6, Appli
23	3411	57.0	1137	10	US-09-902-481A-5	Sequence 5, Appli
24	3401	56.8	1137	10	US-09-902-481A-4	Sequence 4, Appli
25	3388	56.6	1137	10	US-09-902-481A-3	Sequence 3, Appli
26	1833	30.6	494	9	US-09-350-259-103	Sequence 103, App
27	1833	30.6	494	10	US-09-891-943-103	Sequence 103, App
28	1531	25.6	413	10	US-09-350-259-101	Sequence 101, App
29	1531	25.6	413	10	US-09-891-943-101	Sequence 101, App
30	1476	24.7	1170	9	US-09-945-265-2	Sequence 2, Appli
31	1474	24.6	1170	12	US-10-261-164-1	Sequence 1, Appli
32	1462.5	24.4	1223	16	US-10-408-765A-295	Sequence 295, App
33	1318	22.0	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
34	1148	19.2	1179	14	US-10-177-550-2	Sequence 2, Appli
35	1148	19.2	1179	15	US-10-173-551-2	Sequence 2, Appli
36	1146.5	19.1	1189	10	US-09-984-130-35	Sequence 35, Appl
37	1146.5	19.1	1189	10	US-09-836-353A-35	Sequence 35, Appl
38	1146.5	19.1	1189	12	US-10-262-839-4	Sequence 4, Appli
39	1144	19.1	1188	15	US-10-291-265-810	Sequence 810, App
40	1140	19.0	1188	15	US-10-291-265-338	Sequence 338, App
41	1037.5	17.3	1034	10	US-09-984-130-43	Sequence 43, Appl
42	1037.5	17.3	1034	10	US-09-836-353A-43	Sequence 43, Appl
43	1036.5	17.3	1179	12	US-09-918-715-250	Sequence 250, App
44	1030.5	17.2	1151	10	US-09-984-130-103	Sequence 103, App
45	1030.5	17.2	1151	10	US-09-836-353A-103	Sequence 103, App

ALIGNMENTS

```
RESULT 1
US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US2002006200BA1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2002006200BAlel Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-2

Query Match             100.0%;  Score 5987;  DB 9;  Length 1161;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1161;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 TFGTVLLSVLASYHGFNLDVEEPTIFQEDAGFGQSVVQFGGSRLVVGAPLEVVAANQT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TFGTVLLSVLASYHGFNLDVEEPTIFQEDAGFGQSVVQFGGSRLVVGAPLEVVAANQT 60

Qy     61 GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120
```

RESULT 2
US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US20030077278A1
GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-2

Query Match   100.0%;   Score 5987;   DB 10;   Length 1161;
Best Local Similarity 100.0%;   Pred. No. 0;
Matches 1161;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy     1 TFGTVLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TFGTVLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT 60

Qy    61 GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120

Qy   121 GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE 180

Qy   181 GTDTLFALMQYSNLLKIHFTFQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHH 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTDTLFALMQYSNLLKIHFTFQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHH 240

Qy   241 KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN 300

Qy   301 TISSAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TISSAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM 360

Qy   361 DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL 420

Qy   421 GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPH 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPH 480

Qy   481 YYEQTRGGQVSVCPLPRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 YYEQTRGGQVSVCPLPRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI 540

Qy   541 GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD 600

Qy   601 LAVGARGQVLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKS 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 LAVGARGQVLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKS 660

Qy   661 SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCV 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCV 720

Qy   721 EDVVSPILHLNFSLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGV 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 EDVVSPILHLNFSLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDGLCEGDLGV 780

Qy   781 TLSFSGLQTLTVGSSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSA 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TLSFSGLQTLTVGSSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSA 840

Qy   841 LRLACETVPTEDEGLRSSRCSVNEPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSE 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 LRLACETVPTEDEGLRSSRCSVNEPIFHEGSNGTFIVTFDVSYKATLGDRMLMRASASSE 900

Qy   901 NNKASSSKATFQLELPVKIAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQR 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 NNKASSSKATFQLELPVKIAVYTMISRQEESTKYFNFATSDEKKMKEAEHRYRVNNLSQR 960

Qy   961 DLAISINFWVPVLLNGVAVWNDVVNEAPSQSLPCVSERKPPQHSDFLTQISRSPMLDCSIA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 DLAISINFWVPVLLNGVAVWNDVVNEAPSQSLPCVSERKPPQHSDFLTQISRSPMLDCSIA 1020

Qy  1021 DCLQFRCDVPSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVSQLPGQ 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 DCLQFRCDVPSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVSQLPGQ 1080

Qy  1081 EAFMRAQMEMVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLEDKP 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 EAFMRAQMEMVLEEDEVYNAIPIIMGSSVGALLLLALITATLYKLGFFKRHYKEMLEDKP 1140

Qy  1141 EDTATFSGDFSCVAPNVPLS 1161
         ||||||||||||||||||||
Db  1141 EDTATFSGDFSCVAPNVPLS 1161
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2004, 09:37:59 ; Search time 23 Seconds
                                         (without alignments)
                                         2605.989 Million cell updates/sec

Title: US-09-891-943-2
Perfect score: 5987
Sequence: 1 TFGTVLLSVLASYHGFNLD.........DTATFSGDDFSCVAPNVPLS 1161

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued Patents AA:*
              1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
              2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
              3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
              4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
              5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
              6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

              Pred. No. is the number of results predicted by chance to have a
              score greater than or equal to the score of the result being printed,
              and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5987	100.0	1161	3	US-09-193-043-2	Sequence 2, Appli
2	5987	100.0	1161	4	US-09-688-307A-2	Sequence 2, Appli
3	5987	100.0	1161	4	US-09-350-259-2	Sequence 2, Appli
4	5981	99.9	1161	1	US-08-173-497-2	Sequence 2, Appli
5	5981	99.9	1161	1	US-08-286-889-2	Sequence 2, Appli
6	5981	99.9	1161	1	US-08-485-618-2	Sequence 2, Appli
7	5981	99.9	1161	2	US-08-362-652-2	Sequence 2, Appli
8	5981	99.9	1161	2	US-08-605-672-2	Sequence 2, Appli
9	5981	99.9	1161	2	US-08-482-293A-2	Sequence 2, Appli
10	5981	99.9	1161	2	US-08-943-363-2	Sequence 2, Appli
11	5971.5	99.7	1161	1	US-08-485-618-99	Sequence 99, Appl
12	5971.5	99.7	1161	2	US-08-605-672-99	Sequence 99, Appl
13	5971.5	99.7	1161	2	US-08-482-293A-99	Sequence 99, Appl
14	5971.5	99.7	1161	2	US-08-943-363-99	Sequence 99, Appl
15	5971.5	99.7	1161	3	US-09-193-043-99	Sequence 99, Appl
16	5971.5	99.7	1161	4	US-09-688-307A-99	Sequence 99, Appl
17	5971.5	99.7	1161	4	US-09-350-259-99	Sequence 99, Appl
18	4403.5	73.6	1161	3	US-09-193-043-55	Sequence 55, Appl
19	4403.5	73.6	1161	4	US-09-688-307A-55	Sequence 55, Appl
20	4403.5	73.6	1161	4	US-09-350-259-55	Sequence 55, Appl
21	4397.5	73.5	1161	1	US-08-485-618-55	Sequence 55, Appl
22	4397.5	73.5	1161	2	US-08-362-652-55	Sequence 55, Appl
23	4397.5	73.5	1161	2	US-08-605-672-55	Sequence 55, Appl
24	4397.5	73.5	1161	2	US-08-482-293A-55	Sequence 55, Appl
25	4397.5	73.5	1161	2	US-08-943-363-55	Sequence 55, Appl
26	4364.5	72.9	1151	1	US-08-286-889-37	Sequence 37, Appl
27	4364.5	72.9	1151	1	US-08-485-618-37	Sequence 37, Appl

28	4364.5	72.9	1151	1	US-08-362-652-37	Sequence 37, Appl
29	4364.5	72.9	1151	2	US-08-605-672-37	Sequence 37, Appl
30	4364.5	72.9	1151	2	US-08-482-293A-37	Sequence 37, Appl
31	4364.5	72.9	1151	3	US-08-943-363-37	Sequence 37, Appl
32	4364.5	72.9	1151	3	US-09-193-043-37	Sequence 37, Appl
33	4364.5	72.9	1151	4	US-09-688-307A-37	Sequence 37, Appl
34	4364.5	72.9	1151	4	US-09-350-259-37	Sequence 37, Appl
35	4305.5	71.9	1161	1	US-08-485-618-53	Sequence 53, Appl
36	4305.5	71.9	1161	2	US-08-362-652-53	Sequence 53, Appl
37	4305.5	71.9	1161	2	US-08-605-672-53	Sequence 53, Appl
38	4305.5	71.9	1161	2	US-08-482-293A-53	Sequence 53, Appl
39	4305.5	71.9	1161	3	US-08-943-363-53	Sequence 53, Appl
40	4305.5	71.9	1161	3	US-09-193-043-53	Sequence 53, Appl
41	4305.5	71.9	1161	4	US-09-688-307A-53	Sequence 53, Appl
42	4305.5	71.9	1161	4	US-09-350-259-53	Sequence 53, Appl
43	4278.5	71.5	1155	1	US-08-286-889-46	Sequence 46, Appl
44	4278.5	71.5	1155	1	US-08-485-618-46	Sequence 46, Appl
45	4278.5	71.5	1155	1	US-08-362-652-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-09-193-043-2
; Sequence 2, Application US/0919043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-2

Query Match   100.0%; Score 5987; DB 3; Length 1161;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 116; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TFGTVLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TFGTVLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT 60

Qy   61 GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK 120

Qy   121 GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE 180

Qy   181 GTDTLFALMQYSNLLKIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVTQLFHH 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GTDTLFALMQYSNLLKIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVTQLFHH 240

Qy   241 KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQABKAGIIRYAIGVGHAFQGPTARQELN 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQABKAGIIRYAIGVGHAFQGPTARQELN 300

```
;   PRIOR APPLICATION NUMBER: 09/193,043
;   PRIOR APPLICATION FILING DATE: 1998-11-16
;   PRIOR APPLICATION NUMBER: 08/605,672
;   PRIOR APPLICATION FILING DATE: 1996-02-22
;   PRIOR APPLICATION NUMBER: 08/173,497
;   PRIOR APPLICATION FILING DATE: 1993-12-23
;   PRIOR APPLICATION NUMBER: 08/286,889
;   PRIOR APPLICATION FILING DATE: 1994-08-05
;   PRIOR APPLICATION NUMBER: 08/362,652
;   PRIOR APPLICATION FILING DATE: 1994-12-21
;   PRIOR APPLICATION NUMBER: 08/943,363
;   PRIOR APPLICATION FILING DATE: 1997-10-03
;   NUMBER OF SEQ ID NOS: 114
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 2
;     LENGTH: 1161
;     TYPE: PRT
;     ORGANISM: Homo sapiens
US-09-688-307A-2

  Query Match             100.0%;  Score 5987;  DB 4;  Length 1161;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  1161;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TFGTVLLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TFGTVLLLSVLASYHGFNLDVEEPTIFQEDAGGFGQSVVQFGGSRLVVGAPLEVVAANQT   60

Qy     61 GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GRLYDCAAATGMCQPIPLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSK  120

Qy    121 GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GSCLLLGSRWEIIQTVPDATPECPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQFE  180

Qy    181 GTDTLFALMQYSNLLKIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHH  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTDTLFALMQYSNLLKIHFTFTQFRTSPSQQSLVDPIVQLKGLTFTATGILTVVTQLFHH  240

Qy    241 KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 KNGARKSAKKILIVITDGQKYKDPLEYSDVIPQAEKAGIIRYAIGVGHAFQGPTARQELN  300

Qy    301 TISSAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TISSAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  360

Qy    361 DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL  420

Qy    421 GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPH  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPH  480

Qy    481 YYEQTRGGQVSVCPLPRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 YYEQTRGGQVSVCPLPRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI  540

Qy    541 GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD  600

Qy    601 LAVGARGQVLLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKS  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 LAVGARGQVLLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKS  660

Qy    661 SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCV  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCV  720

Qy    721 EDVVSPIILHLNFSLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDLCEGDLGV  780
```

```
Qy    301 TISSAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TISSAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEGTQSRASSSFQHEMSQEGFSTALTM  360

Qy    361 DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 DGLFLGAVGSFSWSGGAFLYPPNMSPTFINMSQENVDMRDSYLGYSTELALWKGVQNLVL  420

Qy    421 GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPH  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAPRYQHTGKAVIFTQVSRQWRKKAEVTGTQIGSYFGASLCSVDVDSDGSTDLILIGAPH  480

Qy    481 YYEQTRGGQVSVCPLPRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 YYEQTRGGQVSVCPLPRGQRVQWQCDAVLRGEQGHPWGRFGAALTVLGDVNEDKLIDVAI  540

Qy    541 GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAPGEQENRGAVYLFHGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMD  600

Qy    601 LAVGARGQVLLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKS  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 LAVGARGQVLLLRSLPVLKVGVAMRFSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKS  660

Qy    661 SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCV  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 SLDQLGDIQSSVRFDLALDPGRLTSRAIFNETKNPTLTRRKTLGLGIHCETLKLLLPDCV  720

Qy    721 EDVVSPIILHLNFSLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDLCEGDLGV  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 EDVVSPIILHLNFSLVREPIPSPQNLRPVLAVGSQDLFTASLPFEKNCGQDLCEGDLGV  780

Qy    781 TLSFSGLQTLTVGSSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TLSFSGLQTLTVGSSLELNVIVTVWNAGEDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSA  840

Qy    841 LRLACETVPTEDEGLRSSRCSVNHPIFHEGSNGTFIVTFDVSVKATLGDRMLMRASASSE  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 LRLACETVPTEDEGLRSSRCSVNHPIFHEGSNGTFIVTFDVSVKATLGDRMLMRASASSE  900

Qy    901 NNKASSSKATFQLELPVKVAVYTMISRQESTKYFNFATSDEKKMKEAEHRYRVNNLSQR  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 NNKASSSKATFQLELPVKVAVYTMISRQESTKYFNFATSDEKKMKEAEHRYRVNNLSQR  960

Qy    961 DLAISINFWVPVLLNGVAVWDVVMEAPSQSLPCVSERKPQHSDFLTQISRSPMLDCSIA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 DLAISINFWVPVLLNGVAVWDVVMEAPSQSLPCVSERKPQHSDFLTQISRSPMLDCSIA 1020

Qy   1021 DCLQFRCDVPSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQ 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 DCLQFRCDVPSFSVQEELDFTLKGNLSFGWVRETLQKKVLVVSVAEITFDTSVYSQLPGQ 1080

Qy   1081 EAFMRAQMEMVLEEDEVYNAIPIIMGSSVGALLLALTTATLYKLGFFKRHYKEMLEDKP 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 EAFMRAQMEMVLEEDEVYNAIPIIMGSSVGALLLALTTATLYKLGFFKRHYKEMLEDKP 1140

Qy   1141 EDTATFSGDDFSCVAPNVPLS 1161
          |||||||||||||||||||||
Db   1141 EDTATFSGDDFSCVAPNVPLS 1161
```